Volume 7 AUGUST, 1936 Number 8

PHYSICS

A Journal of General and Applied Physics

The Thermionic Amplification of Direct Currents

P. A. MACDONALD, *University of Manitoba, Winnipeg, Canada*

(Received May 14, 1936)

The principles and practical applications of thermionic tubes are discussed at some length in relation to the amplification of direct currents.

WITHIN the past six years the use of thermionic valves for the amplification of direct currents, originating in high resistance circuits, has rapidly increased. Various circuits and special valves have been developed so that at present it is possible to measure with precision direct currents of almost any magnitude. It is the purpose of this communication to discuss the principles involved in such measurements and to give specific circuits for various current ranges. The literature on the subject is rapidly becoming voluminous and in the interest of brevity reference will only be made to such publications as seem best to embody definite advances. There are certain basic physical phenomena that must be clearly understood in order that valves may be intelligently employed and attention is first directed to this branch of the subject.

BASIC PHYSICAL PHENOMENA

Electric current

In the common forms of electric circuits the behaviour of the current is almost invariably deduced from the heat generated by its passage or from the magnetic field which it causes about the conductor. The fact that the current is a flow of charges within the conductor is of secondary interest relative to these manifestations. When circuits containing thermionic elements are being treated the situation is altered and attention must be devoted primarily to the atomic nature of the current, employing an outlook similar to that involved when considering electrolytic phenomena. At the outset then, an electric current will be defined as a flow of discrete elementary charges and the number passing any point in unit time defined as a measure of its magnitude. Such charges may be either electrons or positively charged ions and the direction of current flow is defined as being opposite to the direction of motion of the negative ions.

On this basis the presence of a metallic conductor is not a necessary condition for the existence of a flow of current and when such a conductor is employed its function is to be considered that of supplying a large number of free ions and a directive force for their flow. The directive force arises from the existence at the periphery of the wire of a barrier of potential over which the charges cannot pass when their energy is solely derived from the heat content of the conductor at room temperature. This heat energy has its origin in the rapid random motion of the molecular constituents of the wire and introduces complexities into the motion of the ions. These ions, as a result, oscillate irregularly with a mean velocity large relative to their linear velocity resulting from any electromotive force applied to the extremities of the conductor. The kinetic energy of a charge at any one moment is, accordingly, like that of a molecule, a matter of chance. This energy may be increased to a level sufficient to overcome the potential

barrier at the surface of the conductor by raising its temperature and there results a random emission of the charges into space.

If the space surrounding the conductor be evacuated such thermionic emission is similar in behaviour to the evaporation of a gas from a fluid and it may be shown both by theory and experiment that the charges under such conditions have a velocity distribution corresponding to an ionic gas at the temperature of the emitter.[1] Accordingly then in a thermionic vacuum tube the ions are emitted at random and have a random velocity originally derived from the motion of the constituents of the filament, in addition to any velocity arising from an e.m.f. applied across the vacuum.

The random thermal emission of charges with time has been treated by Schottky and others, who calculate the resultant effect upon the space current. Schottky's original paper[2] points out that if discrete charges were emitted in a regular manner, the space current would fluctuate with a frequency measured by the number of charges emitted in unit time and that when the emissions occur at random a fluctuating current of variable frequency is obtained. The so-called steady direct current is only the average value of the fluctuating current measured over a long period of time. By application of the principles of probability and by making the assumption that the random motion of the charges at the terminal electrode in the vacuum is the same as that at the emitter, it is possible to compute the magnitude of the fluctuating current in terms of the long time average or steady current. The mean square value for this quantity over time interval ΔT is

$$\overline{i_{\Delta T}^2} = e i_0 / \Delta T, \tag{1}$$

where e is the ionic charge and i_0 the long time average current. Accordingly then no such thing as a steady space current exists no matter how constantly the conditions governing its flow are maintained. It is however convenient to speak of a steady current and this then has reference to the long time average of the fluctuating current. It is to be noted that the fluctuating component increases with the value of the long time average current.

[1] L. H. Germer, Phys. Rev. **25**, 795 (1925).
[2] W. Schottky, Ann. d. Physik **57**, 549 (1918).

Resistance

The current flowing in a metallic circuit is proportional to the potential difference across the circuit and falls to zero when the e.m.f. is removed.

Stated algebraically

$$i = kE \quad \text{or} \quad E = ri, \tag{2}$$

where i and E are the current and voltage, the constants k and r are defined respectively as the conductance and the resistance of the circuit. In addition to enabling the current to be computed from the voltage the quantity r is of importance in expressing the energy J dissipated in the circuit according to the relation

$$J = i^2 r. \tag{3}$$

If the current be increased from i an amount di then the additional heat liberated from the circuit will be given by

$$\frac{\partial J}{\partial i} di = \frac{\partial}{\partial i}(i^2 r) di = (E + ir) di. \tag{4}$$

Conductors in which these three equations are applicable will in future be referred to as ohmic conductors.

The currents flowing between electrodes in an evacuated space are governed by laws of conduction which are not expressed quite so simply. In general, the current is not proportional to the applied voltage and some current may flow when apparently no e.m.f. is being applied to the circuit. A graphical relation between space current and applied e.m.f. is reproduced in Fig. 1, curve A, which curve has an equation of the form

$$i = \alpha E^{\beta}. \tag{5}$$

By taking logarithms and differentiating,

$$\frac{di}{i} = \beta \frac{dE}{E} \quad \text{or} \quad i = \frac{1}{\beta} \cdot \frac{di}{dE} \cdot E. \tag{6}$$

The quantity di/dE fulfills the same function in Eq. (6) as k in Eq. (2), and is defined as the variational conductance. The reciprocal of this quantity is known as the variational resistance or, more loosely, but more briefly, the impedance. The two quantities, variational conductance and

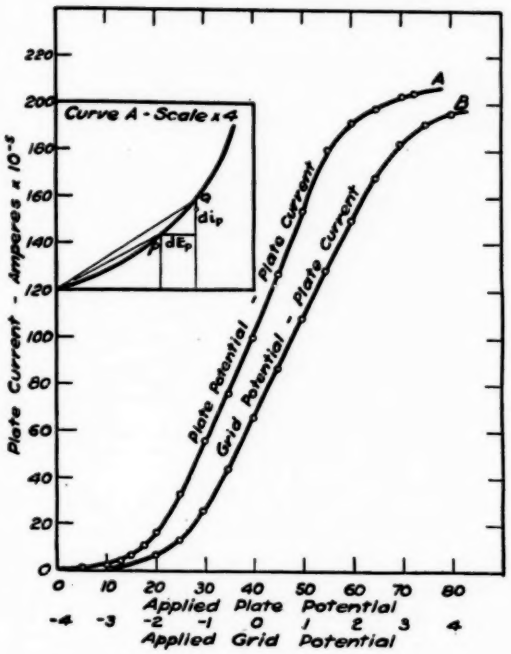

FIG. 1. Curve A shows a space current as a function of the potential applied between an emitting electrode and a collecting plate assembled in a vacuum. Curve B shows the dependence of this curve upon the potential applied to a wire grid placed between the emitter and plate. It will be seen that the grid potential is ten times more effective than the plate potential in controlling the space current. This ratio varies for different vacuum tubes and is defined as the amplification factor of the tube.

variational resistance, will be denoted by K and Z, respectively.

The number of ions carried across a space circuit per second when operated at the point P (curve A) is i, so that the total energy J transmitted from the emitting to the receiving electrode in unit time is given by $J=iE$. This energy reaches the terminal electrode in the form of ionic kinetic energy and is accordingly dissipated as heat in this electrode. Substituting ri for E in this equation yields $J=i^2r$. Accordingly then the quantity r defined by Ohm's law may be employed in computing the total energy dissipated in a space circuit and for this reason it is known as the direct current resistance of such a circuit.

Referring again to curve A the additional heat dissipated by increasing the current from i to an amount $i+di$ and thus going from P to Q on the curve is

$$dJ=(\partial/\partial i)(i^2r)di=(Edi+idE)di$$
$$=(E+idE/di)di=(E+iZ)di. \quad (7)$$

Comparison of Eqs. (2) and (6) and comparison of (4) and (7) demonstrate the relative parts played by the quantities r and Z in metallic and space circuits. The use of the variational resistance is necessitated in space circuits by the non-linear response of current to voltage.

Circuits containing ohmic resistors have associated with them certain phenomena resulting from the random motion of the electric charges in the resistor. This motion causes a fluctuating e.m.f. across the terminals of the resistor. The possibility of the existence of such an e.m.f. was first suggested by Schottky[2] and discovered by Johnson.[3,4] Nyquist[5] theoretically evolved from thermodynamic considerations the following expression for its magnitude

$$E^2dn=4K_bTrdn, \quad (8)$$

where E^2dn is the square of the e.m.f. generated within the frequency range dn, K_b is Boltzmann's constant, T is the temperature of the resistor in degrees absolute and r is the value of the resistor. In deriving this formula absolutely no restriction has been placed upon the material employed to construct the resistor; accordingly whenever an ohmic resistance is placed in a circuit the presence of such a fluctuating e.m.f. must be taken into consideration and the above equation may be used to compute its value.

The resistance associated with the passage of a space current has also been shown to give rise to a fluctuating thermal e.m.f. Llewellyn[6] on a theoretical basis concludes that Eq. (8) may be used to compute its value when T is given the value of the temperature of the emitter. In addition to this thermal e.m.f. other sources of variation resulting from resistors are often introduced into a circuit when it is not feasible to construct them of wire. A practical upper limit is placed upon the magnitude of a wire wound resistor by its dimensions and cost. For values in excess of 10^7 ohms it is customary to employ materials other than wire and when the resistors in a circuit are made extremely great those parts of the circuit primarily intended for insulators must also be considered as con-

[3] J. B. Johnson, Phys. Rev. **29**, 367 (1927).
[4] J. B. Johnson, Phys. Rev. **32**, 97 (1928).
[5] H. Nyquist, Phys. Rev. **32**, 110 (1928).
[6] F. B. Llewellyn, Proc. Inst. Rad. Eng. **18**, 248 (1930).

ductors; accordingly the electrical properties of even insulating glasses and quartz must be given special study.

Some special phenomena of resistance are encountered for extremely high values and may be classified under the heading of electrolytic polarization and space charge. Electrolytic polarization is found to exist in many types of high resistors and the building up of a space charge in glass with an applied voltage has been the subject of considerable investigation.[7, 13] Both these phenomena give rise to the same effect, namely, the variation of the value of a high resistance unit with time and with the magnitude of the impressed e.m.f.

In conclusion then, wherever resistance is present in a circuit allowance must be made for the existence of a fluctuating thermal e.m.f. across its extremities and further when such resistance is high a variation of its value with time and with the magnitude of the impressed e.m.f. may be expected.

Thermionic Valve

A thermionic valve consists essentially of a space evacuated of as much gas as possible and containing an emitter of ions, a collector or plate and a grid system.

The grid is of wire placed between the emitter and plate in such a manner that any electrostatic field established between them will be disturbed by a potential applied to the grid.

The standard diagram for such a three element tube or triode and its associated circuit is given in Fig. 2, consideration of which makes it possible to determine the function of the tube as an amplifier. With reference to Fig. 2A the emitter E_m, heated by the battery A, liberates both positive ions and electrons. Under many conditions the emitted electrons exceed the positive ions in number by a factor of the order of 10^8. A current is maintained between the emitter and plate P by the battery B, which charges the plate positively. The current is accordingly electronic, provided the residual gas in the tube is not ionized. The current as a function of the potential due to the battery B has been reproduced graphically in Fig. 1, curve A.

[7] D. M. Robinson, Physics **2**, 52 (1932).

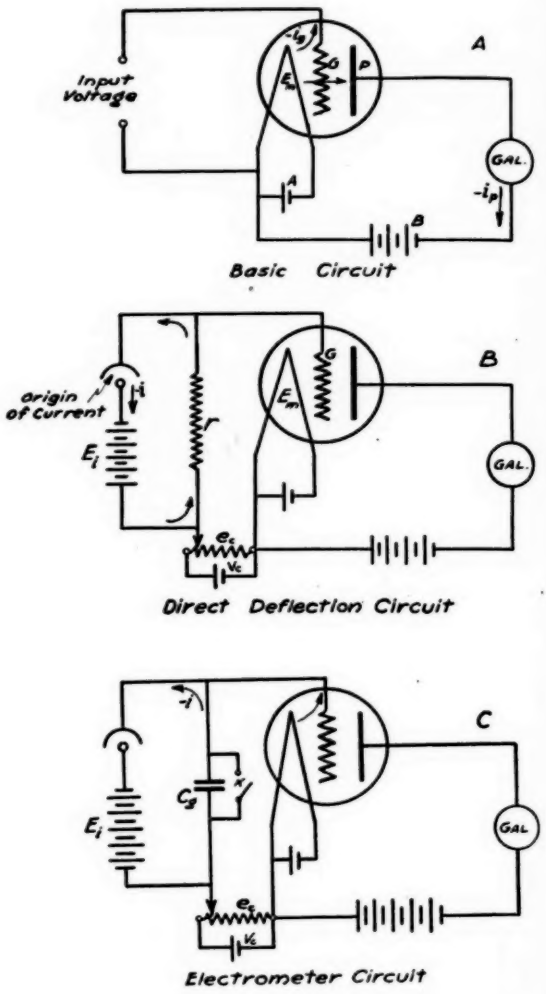

Fig. 2. Fig. 2A shows how the amplification factor of a tube is employed to amplify a voltage. The input voltage is impressed between the grid and filament and the corresponding change in plate current observed. In Fig. 2B the photoelectric current is used to produce the voltage by forcing the current through the high resistance r, the resulting e.m.f. being measured as in Fig. 2A. In Fig. 2C the vacuum tube shows the rate of accumulation of charge on the condenser C_g by the rate of change of the plate current. The value of the photoelectric current is computed from the rate of accumulation of the condenser charge.

The fact that the space current is dependent upon the applied potential shows that there exists an opposing e.m.f. elsewhere in the circuit. This opposing e.m.f. has its origin in the cloud of electrons between the filament and plate, such a space charge hindering the passage of electrons from the filament toward the cloud and hence toward the plate. Curve B in the same figure shows the effect of the grid potential upon the

plate current and the form of this curve emphasizes the fact that the grid potential only exercises control over the plate current in the presence of a space charge. The value of a thermionic valve as a measuring instrument lies in this ability of the potential of the grid to control the plate current in spite of the fact that it is not connected in the plate circuit.

A valve itself is then fundamentally a voltage operated instrument and any current to be measured by such a device must first be made to produce a corresponding voltage. Such a transformation may be brought about in two ways. Either the current in question may be forced through a high resistance (r, Fig. 2B), the resulting drop in potential being impressed between the filament and grid of the tube, or the grid-filament system may be treated as a condenser (Fig. 2C). The grid side of this condenser is charged up by forcing the current on to it and the rate of change of the potential of the system measured by observing the corresponding rate of change of the plate current. The former type of amplification is termed the direct deflection method, the latter instrument is known as the vacuum tube electrometer. This simple description only represents the action of an amplifier to a first approximation owing to the existence of a space current which flows between the filament and the grid. This grid current introduces a high degree of complexity into what would otherwise be a simple instrument and places a lower limit upon the current which may be measured. Employing a valve for amplifying currents accordingly involves the use of three distinct circuits, first the circuit in which the current originates, secondly the filament-grid circuit of the valve and finally the filament-plate circuit of this device. Of the currents flowing in these three circuits that of the grid is the most complex and is of such importance as to warrant detailed study.

Grid circuit

This circuit, which in Fig. 2B may be traced as $E_m G r e_c$, includes the emitter and grid within the evacuated tube, an external or input resistance and a source of potential e_c, the manipulation of which enables the flow of grid current to be controlled.

Attention is first directed to the properties of the emitter. The only type which at present is suitable for use with direct current amplifiers is a heated filament of wire. Other sources of emission are either too irregular in their action as in "heater tubes" used to operate alternating current radios or are incapable of supplying ions in sufficient quantities to be of practical value. In future, then, the emitter will be referred to as the filament.

It has been pointed out previously that the number of ions emitted is dependent upon their velocity within the conductor, which velocity is a function of its temperature. The maximum number emitted at any temperature T is given by the equation $i = AT^2\epsilon^{-W_0/K_bT}$, where W_0 is the work function at absolute zero, K_b Boltzmann's gas constant for a single electron and A a constant, which theoretically has the same numerical value for all pure metals under specified conditions. Practical examination confirms this theoretical conclusion. However, experiment shows, that if the emitting metal be coated with an element more electro-positive, the value of A is reduced: coating with a more electro-negative element on the other hand increases A. Accordingly it is possible to choose an element having certain desirable physical characteristics for the construction of a filament and adjust its emissive power by suitable treatment with other elements. The following data show the effect of coating tungsten with the more electro-positive thorium.

Constant	Pure Tungsten	Thoriated Tungsten
A	60.2	3.0
W_0/K_b	52,400	30,500

Tungsten is commonly used for filaments and has been the subject of intensive study. The great objection to the use of this metal is its low emissive power; however, practically all tubes suitable for direct current work operate with a tungsten filament or tungsten coated with the more electro-positive thorium, such complex emitters being known as thoriated tungsten filaments.

Thorium has a melting point of 2118°K in contrast to 3655°K for tungsten, but a film of thorium upon a tungsten surface is prevented from evaporating readily by some surface

phenomenon and the operating temperature of the filament may be as high as 2000°.[8] One characteristic of amplifiers employing thoriated filaments is a slow steady drift of the recording instrument attributable to distillation of thorium.

All the properties of tungsten that are of interest when this metal is employed for radio tube filaments have been discussed in great detail by Jones and Langmuir.[9] In addition to these properties the emission of positive ions is of extreme importance when steady currents are to be amplified. The subject of positive ion emission has been discussed by Smith[10] both for tungsten and molybdenum. He found that for temperatures in excess of 2300°K both metals emit their own singly charged positive ions but no doubly charged ions or single charges with a double mass. The emission from tungsten was found to depend upon previous heat treatment, and even after ageing there was found a slow decay in emissivity with time. Molybdenum on the other hand was found to give constant emission.

Smith further showed that the form of the work function for positive ions is quite different from that of electrons but that they are related according to the equation

$$W^{+}_{T}+W^{-}_{T}=V+U_{T},$$

where W^{+}_{T} and W^{-}_{T} are the positive and negative ion work functions respectively at the temperature T, V is the potential necessary to form the positive ion and U_T is the heat of evaporation.

This relation between the work functions shows that impurities upon a filament (e.g. thorium) will materially affect the degree of positive ion as well as the degree of electronic emission.

In addition to the positive ions of the metals examined, Smith found ions of sodium, potassium and aluminum emitted at temperatures of 1700°K, 1700°K and 2000°K, respectively. The impurities recognized in this manner were not present in sufficient quantities to be detected spectroscopically. It was found that heating the filament to a high temperature rapidly removed the positive ions of the foreign metals which would allow one to conclude that it should be possible to operate a pure tungsten filament at a temperature below 2300°K without the emission of any positive ions.

[8] L. R. Koller, *Physics of Electron Tubes* (McGraw-Hill, 1934).
[9] J. A. Jones and L. Langmuir, Gen. Elec. Rev. **30**, 310 (1927).
[10] L. P. Smith, Phys. Rev. **35**, 381 (1930).

Control grid

In addition to the ions emitted from the filament, radiation of heat and light from this source act upon the grid so that it becomes a thermionic and photoelectric emitter. Electrons emitted from the grid by these agencies have the same end result on the grid current as increasing the positive ion emission from the filament. The rate of emission of light from the filament as a function of the filament temperature is shown in Fig. 3. In addition to thermionic and photoelectric emission, x-ray photoelectrons are also emitted from the grid, the x-rays having their origin at the plate as a result of the impact of the plate current electrons.

In Fig. 4 is reproduced graphically the current flowing in the grid circuit as a function of the applied grid potential, which current is the summation of the constituents enumerated above. With potentials positive relative to the filament the curve is similar to curve B, Fig. 1, relating the plate potential to the plate current. This part of the curve is of no particular interest

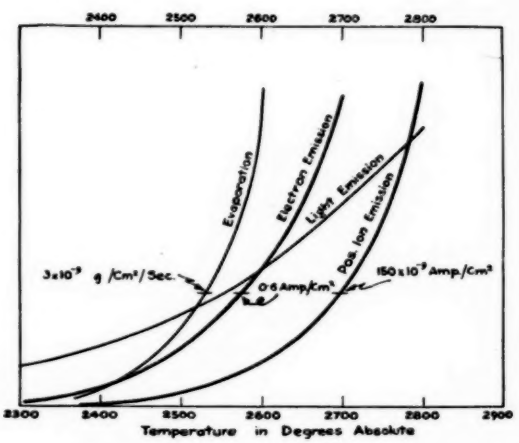

FIG. 3. Emission from tungsten as a function of the temperature. These various types of emission complicate the grid current. Since the electron emission curve dies off less rapidly than the other curves the best condition for a pure electron grid current is obtained with a low filament temperature.

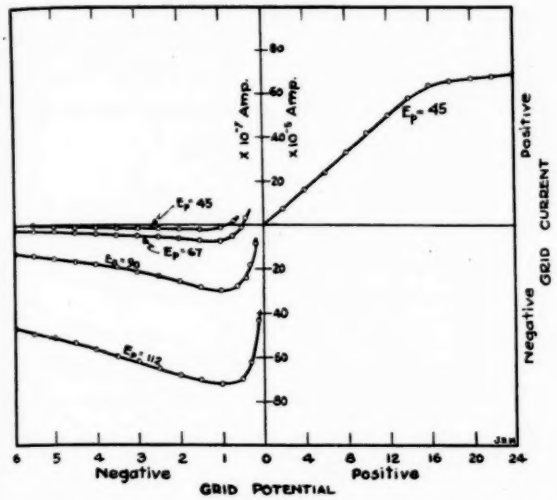

FIG. 4. The family of curves on the left shows how the grid current increases with increase of the applied plate potential E_p. This is due to the increased formation of gaseous ions at higher potentials. Direct current amplifiers are invariably operated with a negative potential on the grid in order to keep the energy in the grid circuit a minimum. The right hand curve shows the electronic grid current and is plotted on a scale 100 times smaller than the curves on the left.

in the study of direct current amplifiers owing to the power consumed, for as has already been pointed out, the value of the tube lies in the electrostatic control of the plate current by the grid potential. The grid current flowing with a negative potential on the grid is on the other hand of extreme importance in spite of its magnitude. This section of the curve has been discussed in some detail by Nottingham[11] in 1930 and the subject has subsequently been given further study in the laboratories of the General Electric Company.[12, 13]

With negative potentials on the grid the electronic flow may be completely stopped if desired or it may be reduced to the same order of magnitude as the positive ion current. One of the most important potentials of the grid is that at which these two currents are numerically equal so that over a period of time there is no resultant transfer of charge between filament and grid. This point is defined as floating grid potential since it is this value that an insulated grid attains to on reaching equilibrium. A tube is almost invariably operated somewhere in the vicinity of the floating grid potential so that each of the small constituents of the grid current may be of such a magnitude relative to the whole to be of major importance in determining the behavior of the grid current. The manner of the exact control of these constituents is now taken up.

[11] W. B. Nottingham, J. Frank. Inst. **209**, 287 (1930).
[12] G. F. Metcalf and B. J. Thomson, Phys. Rev. **36**, 1489 (1930).
[13] A. W. Hull, Physics **2**, 409 (1932).

CONTROL OF GRID CURRENT CONSTITUENTS

Electron grid current

In addition to the control of the electron current by the grid potential as shown in Fig. 4 this current is found to increase with increasing filament emission.

Gaseous positive ions

Even within the most highly evacuated space obtainable it has been shown that a gaseous positive ion current in excess of 10^{-13} amp. exists if the potentials employed across the space are sufficient to cause ionization.[12] With all potentials reduced below 8 volts the gaseous ion current disappears. A study of the variation of this gaseous ion current in a commercial radio tube (222 used as a triode) with the applied plate potential is reproduced in Fig. 4. In general it may be stated that for all except electrometer circuits 12 volts may be used across a tube without introducing any undesirable quantity of gaseous ions.

Metallic positive ions

In order to reduce the flow of metallic positive ions without reducing the filament temperature it is customary to employ a second grid placed between the filament and control grid and apply to this grid a potential positive relative to the filament. This additional grid is known as a space charge grid. The information given by Smith does not yet seem to have been successfully incorporated in a vacuum tube for the elimination of positive ions. The indications are, however, that a pure tungsten filament would be superior to a thoriated emitter as it will later be shown that high electronic emission is not always necessary for the successful operation of d.c. amplifiers.

Photoelectronic current

The magnitude of the photoelectric current from the grid is so large that a sensitive ampli-

fier will not operate when exposed to daylight. Light from the filament is also a source of trouble and the present method of reducing it to a minimum is to employ a thoriated tungsten filament, so that the lower operating temperature involved both decreases the light intensity and shifts the wave-length toward the red end of the spectrum. With certain tubes especially constructed to yield low grid currents, the control grid is made of molybdenum with nickel supports since both of these metals have a high photoelectric threshold. Unfortunately the use of a thoriated filament results in a deposit of thorium forming upon this grid which alters the value of its photoelectric threshold. It has been stated[13] that, at the operating temperature commonly employed for such tubes, the photoelectric emission from the grid could probably be reduced by a factor of 10 with the removal of the deposited thorium. In addition to electrons liberated by the action of light from the filament the grid is exposed to the action of the x-rays from the plate. The magnitude of the grid current resulting from this source increases with increasing plate potential and also increases with the lowering of the photoelectric threshold of the grid, due to thorium contamination. The x-ray effect upon a grid contaminated with thorium is detectable with a plate potential of 2.65 volts, which is the lowest photoelectric threshold of thorium.

In summing up the methods of controlling the constituents of the grid current it may be said that in general the factors that are required to give a large plate current increase the grid current. Thus when an extremely small grid current is required attention should be devoted to decreasing the demands placed upon the plate current.

Analytical Considerations

The relation between the grid potential and the grid current has been shown graphically in Fig. 4, the relation between the grid potential and the plate current in Fig. 1. It is possible to correlate analytically these data, together with those relating the input current to the grid current and deduce a general equation governing the behavior of all amplifiers.

The design of an amplifying tube is such that a change in the potential of the grid will cause a change in the plate current greater than would be caused by the same change in the plate potential.

Stated algebraically,

$$di_p/dE_g = \mu di_p/dE_p. \quad \mu > 1 \tag{9}$$

The factor μ is defined as the voltage amplification factor and often arises in the form:

$$\mu = dE_p/dE_g. \tag{10}$$

Following the ideas of resistance and conductance for space currents previously discussed, the expression di_p/dE_p is defined as the plate variational conductance and will be denoted by K_p.

$$K_p = di_p/dE_p. \tag{11}$$

The reciprocal of this quantity, Z_p, is defined as the plate variational resistance or plate impedance

$$Z_p = dE_p/di_p. \tag{12}$$

The expression di_p/dE_g is defined as the mutual conductance and will be denoted as K_m.

$$K_m = di_p/dE_g. \tag{13}$$

The relation between K_m and μ is accordingly

$$\mu = K_m Z_p. \tag{14}$$

It is to be emphasized that the mutual conductance as defined above refers only to a property of the tube. When the characteristics of the associated circuit are included the corresponding relation, which may differ greatly in value, is termed the effective mutual conductance and will be referred to as K_{em}.

In order to derive an analytical expression for the effective mutual conductance it is convenient to consider the sources of potential in the plate circuit.

The effective potential E_p between filament and plate may be written

$$E_p = E_B - i_p r_p + \mu E_g + E_j + E_0, \tag{15}$$

where E_B is the potential applied by the plate battery
E_j is a potential due to the kinetic energy of the charges liberated from the filament by the energy of the filament battery
E_0 is the sum of the contact potentials in the circuit
and μE_g is equivalent to a battery placed in the plate circuit.

The plate current accordingly may be written

$$i_p=(E_B-i_pr_p+\mu E_g+E_i+E_0)/Z_p. \quad (16)$$

Differentiating

$$\frac{di_p}{dE_g}=\frac{\mu}{Z_p}-\frac{di_p}{dE_g}\cdot\frac{r_p}{Z_p} \quad (17)$$

or $$di_p/dE_g=(\mu/Z_p)/(1+r_p/Z_p) \quad (18)$$

or $$K_{cm}=K_m/(1+r_p/Z_p). \quad (19)$$

The relation between the input current i and the grid potential E_g is readily obtained from a consideration of Fig. 2B. From the figure

$$E_g=e_c-i_gr+ir. \quad (20)$$

Differentiating

$$dE_g/di=r-r(di_g/di)\cdot dE_g/dE_g \quad (21)$$

$$=r-r(di_g/dE_g)\cdot(dE_g/di) \quad (22)$$

$$=r-K_grdE_g/di, \quad (23)$$

$$dE_g/di=r/(1+rK_g)=1/(k+K_g), \quad (24)$$

where K_g is defined as the grid conductance and k is the conductance of the input resistance r.

The effect of the input current i upon the plate current i_p is given by

$$di_p/di=(dE_g/di)\cdot di_p/dE_g. \quad (25)$$

Substituting (19) and (24) in (25)

$$di_p/di=K_m/(k+K_g)(1+r_pK_p), \quad (26)$$

which equation expresses the overall amplification and contains all the parameters of a three element tube.

Amplifier sensitivities

When amplifiers are primarily used for measurement, rather than counting and operating relays, interest lies not so much in the change in plate current resulting from the input current as in the change of the number of scale divisions deflection of the meter in the plate circuit.

Denoting this number by n, the voltage sensitivity S_v of a circuit may be written

$$S_v=dn/dE_g=(di_p/dE_g)\cdot dn/di_p=K_{cm}S_g \quad (27)$$

when S_g is the sensitivity of the plate current meter expressed in divisions deflection per unit current.

The sensitivity to current accordingly may be written

$$S_i=\frac{dn}{di}=\frac{dn}{dE_g}\cdot\frac{dE_g}{di}=\frac{1}{k+K_g}\cdot K_{cm}S_g \quad (28)$$

or $$S_i=K_mS_g/(k+K_g)(1+r_pK_p), \quad (29)$$

which equation contains all the parameters of an amplifying circuit operated by a three element tube.

Of the six parameters occurring in this equation only the grid conductance will be discussed in detail as the behaviour of the majority of the remainder is either self-evident or more clearly brought out in discussions on specific circuits.

The grid conductance from its definition may be determined numerically by measuring directly the slope of the grid-potential *vs.* grid-current curve, which method is usually employed when the conductance is extremely small. The grid current itself is usually too small to be measured directly with a galvanometer but may be conveniently determined by examining the rate at which a charge accumulates on the grid when this element is isolated from all but the grid current, the rate of change being determined by the rate of change of the deflection of the plate current meter. If C_g denote the capacity of the grid system then

$$i_g=C_gdE_g/dt. \quad (30)$$

If the value of the grid conductance is of the order of 10^{-11} reciprocal ohm or greater it is necessary to add additional capacity across the grid-filament system in order to successfully employ the timing method. A second method, and more convenient to employ under this condition, is to determine the value from two applied grid-potential *vs.* plate-current curves. One such curve is secured with no resistor in the grid circuit, a second curve is then obtained by repeating the first series of measurements after a resistor has been included in the grid circuit. With the second curve the applied grid potential is not the true grid potential since a drop has occurred across the grid resistor. In this instance the grid potential E_g is related to the potential V_g applied to the grid circuit by the equation

$$E_g=V_g-ri_g. \quad (31)$$

Differentiating

$$di_g/dE_g=(1/r)(dV_g/dE_g-1), \tag{32}$$

which may be written

$$K_g=-\frac{1}{r}\left(\frac{dV_g}{di_p}\cdot\frac{di_p}{dE_g}-1\right). \tag{33}$$

The numerical values of dV_g/di_p and di_p/dE_g are determined from the slopes of the curves.

The variation of grid conductance with the operating potentials of a tube follows closely the variation of the grid current save at floating grid potential, so that the information previously given regarding the manner of control of the grid current is immediately applicable to control of the grid conductance.

At floating grid potential, while the current is zero, the conductance is finite and considerably larger than with potentials slightly more negative. The explanation is, of course, that at floating grid potential there are really two grid currents flowing in opposite directions. In concluding these remarks on grid conductance it is to be noted that the shape of the grid-potential *vs.* grid-current curve shows that the grid impedance may have negative as well as positive values.

Thermionic Amplifying Circuits

General limitations

There are four distinct characteristics of thermionic circuits which place a lower limit upon the value of the current that may be measured and which are of basic importance in designing sensitive circuits. These characteristics are the time constant of the circuit, the thermal or Johnson effect arising in the high resistance section of the circuit and two different phenomena having their origin in erratic thermionic emission.

Time constant

The time constant deals with the rapidity of response of the change in grid potential resulting from a change in input current. Referring to Fig. 2B which depicts a photoelectric cell connected in series with a source of potential and the input resistor r of the amplifying system, the grid will begin to charge up positively when the cell is illuminated but will not reach its final value instantaneously. Assuming that the grid current is very much smaller than the input photoelectric current the rate of growth of the newly acquired potential E is given by

$$dE/dt=i/C, \tag{34}$$

where i is the photoelectric current and C is the capacity of the grid system which in this instance consists of the grid to filament capacity, the capacity of the photoelectric cell and that of the associated leads. The rise in potential of the grid tends to cause a second current to flow through the resistor in the opposite direction to the input current so that after a small interval of time the equation for charging becomes

$$dE/dt=(1/C)(i-E/r), \tag{35}$$

which has the solution

$$E_T=ir(1-e^{-T/rC}) \tag{36}$$

$$=E_0(1-e^{-T/rC}), \tag{37}$$

where E_T is the rise in potential that has occurred after a time interval T and E_0 is the total increase in potential that will have been attained when the circuit has reached equilibrium

In order to anticipate the behavior of any specific circuit and for the purpose of comparing different circuits, a time constant rC has been defined. The dimensions of rC are those of a time and the substitution of the numerical value of this quantity in the above equation yields

$$E_{rC}=0.632E_0. \tag{38}$$

This relation states that in any circuit the grid will attain to 0.632 of its final potential after the input current has been applied over the time interval obtained by multiplying the capacity of the grid system by the value of the input resistor. After a period five times greater than the time constant the grid potential will have reached approximately 99 percent of its final value. After a period given by $10rC$ the circuit is for all practical purposes stable. The time constant accordingly places a limit upon the values of r and C that may be employed. The capacity of the grid-filament system of a thermionic tube lies between 3×10^{-12} farad and 5×10^{-12}

farad. A spherical photoelectric cell may have approximately the same capacity. The sum of these capacities together with those of the leads shows that it is hardly possible to operate with a total capacity much less than 10^{-11} farad. Employing a resistor of 10^{11} ohms will then yield a time constant of one second so that approximately ten seconds will be required for each determination of the input current. A knowledge of the time constant emphasizes the value that is to be placed upon keeping the capacity of the grid system a minimum. In practice the upper limit of input resistance that may be successfully employed seldom exceeds 10^{11} ohms.

Shot effect

It has already been stated that a thermionic current is equivalent to a true steady current upon which is superimposed a small fluctuating component. Provided the fluctuating component is not distorted by other manifestations of space conduction, such as a space charge, it has further been shown that the relation for the fluctuating component is given by

$$\overline{i_{\Delta T}^2} = ei_0/\Delta T. \tag{1}$$

This fluctuating current in the grid circuit will alter periodically the potential of the control grid and the magnitude of the effect for a direct current amplifier has been computed by Hafstad and Brent[14] on the basis of the original work carried out by Schottky. The desired equation is developed by expressing the space current i_g as the sum of a large number of periodic currents, each component current i_n having an angular velocity n times the fundamental period of the series, and the successive currents differing in frequency by the element of frequency $2\pi/T$. These conditions are expressed analytically by the equation

$$i_g = \sum_0^\infty A_n \sin(\omega_n t + \varphi_n), \tag{39}$$

where $\omega_n = 2\pi n/T$ and $\Delta\omega = 2\pi/T$.

The quantity A_n has been shown by Schottky to have the same value for all the component currents and to be given in value by

$$A_n = (4ie/T)^{\frac{1}{2}}, \tag{40}$$

where T is the time interval over which the current is observed. The e.m.f. generated between the filament and grid as a result of such current passage is the product of the current by the alternating current impedance of the path.

$$\overline{\Delta E_g}^2 = \frac{4ie}{T} \sum_0^\infty \sin^2(\omega_n t + \varphi_n) \frac{Z_g^2}{(1+\omega^2 C_g^2 Z_g^2)}. \tag{41}$$

In order to obtain a definite numerical value for the voltage Brent evaluates this summation by integration and obtains the simple expression

$$\overline{\Delta E_g}^2 = ei_g Z_g/2C, \tag{42}$$

where e is the ionic charge. This equation shows the magnitude of the irregular potential on the grid and hence the magnitude of the irregular fluctuation of the plate current.

Thermal or Johnson effect

The magnitude of this effect has also been computed by Hafstad for a direct-current amplifier[14] employing Nyquist's equation.

The element of voltage produced over the frequency range dn has been given as

$$\overline{\Delta E}^2 = 4r_n K_b T dn. \tag{8}$$

The value of r_n is for the grid circuit

$$r_n = r/(1+\omega^2 C_g^2 r^2). \tag{43}$$

Hence the total voltage fluctuation, resulting from the summation of the voltage for each frequency component, is

$$\overline{E}^2 = \frac{4K_b T}{2\pi} \int_0^\infty \frac{r}{1+\omega^2 C_g^2 r^2} d\omega \tag{44}$$

$$= \frac{K_b T}{C_g}, \tag{45}$$

an expression independent of the value of r.

Flicker effect

In examining the shot effect as it applied to alternating current amplification Johnson[15] has found considerably greater fluctuations than anticipated for low frequencies and has concluded that this results from secular changes in the surface conditions of the emitter. For a tungsten

[14] L. R. Hafstad, Phys. Rev. **44**, 207 (1933).

[15] J. B. Johnson, Phys. Rev. **26**, 71 (1925).

filament at a frequency of ten cycles the observed fluctuation is fifty times in excess of the calculated, while at 200 cycles the ratio of computed to observed is 0.7. With an oxide-coated filament the effect is a great deal more pronounced being at 100 cycles 100 times the computed value. With the coated filament the predicted and observed values for the shot effect agree at 5000 cycles.

In considering the practical effects resulting from the existence of atomic electrical charges the foregoing equations are of value insofar as they indicate the manner in which the irregularities arise and show the parameters which are available for investigation of the phenomena.

A careful study of the resultant fluctuations in the plate circuit has been made by Hafstad with FP–54 tubes operated at floating grid potential. He has found a fluctuation corresponding to a grid potential change of 3×10^{-5} volt. The present author has also obtained similar results for both FP–54 and UX–222 tubes. A study of the literature describing amplifiers employed for various purposes indicates that all precisely constructed circuits exhibited fluctuations of the same order of magnitude. The numerical value of the fluctuation obtained for the shot effect from Eq. (42) is 7×10^{-5} volt and for the Johnson effect it is 3×10^{-5} volt. The observed values agree with those computed for the Johnson effect. The large value computed by the shot equation may be due to the manner of deriving the formula, incorporating as it does the assumption that the random emission from the emitter reaches the collecting electrode without distortion. The present state of knowledge of these phenomena in relation to direct current amplification is unsatisfactory. It may, however, be taken as empirically established that a sensitivity of 3×10^{-5} volt represents the maximum voltage sensitivity that may be realized for an amplifier operated at room temperature without increasing the capacity of the grid system and that this fluctuation has its origin in the atomic nature of the electric charge. The capacity of the grid system is invariably increased when the source of input current is connected to the grid and this will allow the voltage sensitivity to be increased somewhat, but not the current sensitivity, since the input resistor must be decreased in order to readjust the time constant.

Tube characteristics and basic circuit

Tubes at present available and suitable for direct current amplification fall into three main classes; triodes, which contain only the three essential elements, filament, control grid and plate; screen grid tubes which contain a second grid placed between the control grid and the plate and finally space charge grid tubes which also contain a second grid but which in this instance is placed between the filament and the control grid. In addition to the class of tube there is a choice of thoriated tungsten or pure tungsten filaments. The type of tube employed depends to a considerable extent upon the type of work to be carried out. Amplification has one of two ends in view, either to measure currents too small to be recorded directly by a galvanometer or to supply sufficient energy to operate some type of relay or recording instrument.

Considerations are at present limited to amplifiers employed for the first class of work and their design is largely controlled by the properties of the galvanometer to be used in the plate circuit. Rapidity of response and ability to stand up to sudden overloads are characteristics that should be maintained to the greatest possible extent. The most satisfactory type of instrument that the author has used is the type R, Leeds and Northrup reflecting galvanometer. Suspensions for these galvanometers are readily mounted and different types of coils are available for the same magnetic system. Utilizing interchangeable coils the following characteristics are obtainable in one instrument,

Type	A	B	C	D
Sensitivity	10^{-10} amp./mm	5×10^{-10}	5×10^{-9}	3×10^{-9}
Period	14	6	5	3
Coil Resistance	500	500	40	500
Critical Damping Resistance	22,000	10,000	300	2,000

All instruments but type C are excellent. The low critical damping resistance of the latter renders its use less desirable than the others for reasons that will subsequently appear. Type D is practically as satisfactory as a needle pointer instrument. The sensitivity of these instruments is expressed for a scale at one meter distance.

This sensitivity, however, may readily be increased as much as five times by employing good optical equipment, as that supplied by the gal-

vanometer manufacturers is capable of considerable improvement. If a lamp and scale are to be employed an achromatic lense of some three centimeters in diameter with a focal length of from 6 to 10 cm, focussed upon a fine hair illuminated by a motor car head light bulb, will give excellent definition at a distance of from 2 to 3 meters from the galvanometer mirror. A most satisfactory type of telescope may be obtained from an engineer's surveying level equipped with a high power eyepiece and spiderweb crosshairs. Galvanometers more sensitive than the foregoing are on the market but their use is not usually desirable as the period is long, the suspension delicate and its replacement often difficult

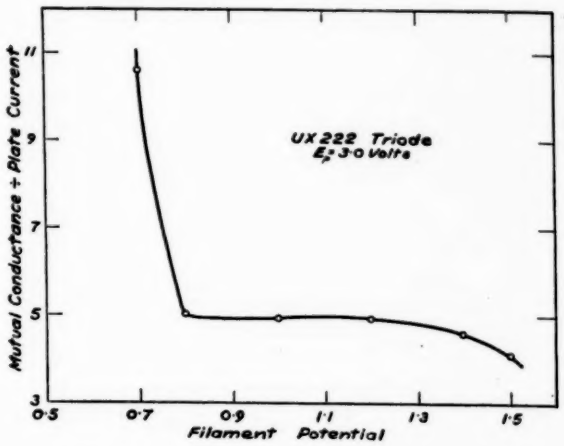

FIG. 5. Lowering the filament potential decreases the mutual conductance less rapidly than the plate current. Circuits in which the magnitude of the plate current governs the galvanometer sensitivity are accordingly more sensitive when the filament potential is kept low as a more sensitive galvanometer may then be employed.

Basic circuit

All circuits are adaptations of the fundamental circuit shown in its simplest form in Fig. 2.

The sensitivity of this system to voltage S_v may be written immediately from the figure as

$$S_v = K_{em} S_g, \tag{46}$$

where S_v is expressed in mm/volt and S_g, the meter sensitivity in mm/amp. The sensitivity to current S_i is accordingly

$$S_i = r K_{em} S_g. \tag{47}$$

Various types of complex circuits employed are extensions of this basic circuit and devised for the purpose of controlling the values of the parameters of the basic equations. Attention usually is directed to increasing S_g by decreasing the initial current flowing through the galvanometer and thus allowing a more sensitive instrument to be employed. By the initial current is meant that which flows when the grid is not being excited by the input current.

The basic circuit is itself capable of yielding highly satisfactory results with a current sensitivity as high as 10^{-15} amp./mm employing suitable radio tubes operated to yield special characteristics. In this simple circuit the sensitivity of the galvanometer that may be employed is limited by the magnitude of the initial plate current all of which passes through this instrument. The plate current is a function of the operating potentials of the valve, and in general decreases as the operating potentials are decreased. K_{em} is also a function of these potentials, also generally decreasing with decreased potentials. The maximum voltage sensitivity of the basic circuit will result when the operating potentials are such that the quotient of the mutual conductance by the plate current is a maximum. This quotient as a function of the filament potential is reproduced in Fig. 5, for one specific plate potential. It is not desirable to utilize the high values resulting from reducing the filament potential below 0.8 volt because of the high slope of the curve in this region, for this introduces instability in the circuit by registering small fluctuations in the filament supply battery. On this particular curve values between 0.8 and 1.2 volts will be the most satisfactory as in this region the slope is a minimum and the value of the quotient greater than for all higher potentials.

The quotient as a function of the plate potential and determined for the most satisfactory filament potential at each plate voltage, is reproduced in Fig. 6. This curve is typical of practically all tubes.

The information gained from these curves shows that the lower the applied plate potential the greater the sensitivity of the unit save with zero volts applied to the plate in which case the sensitivity drops off slightly. With the particular tube used to determine the foregoing data the voltage sensitivity was found to rise from 3×10^{-2} volt/mm at 90 volts on the plate to 5×10^{-4}

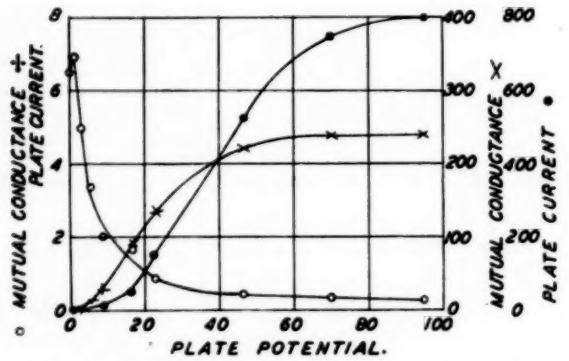

FIG. 6. Data showing, after the manner of Fig. 5, how the sensitivity to voltage increases with decreasing plate potential. The mutual conductance divided by the plate current is proportional to the sensitivity measured in millimeters per volt.

volt/mm with a plate potential of 1.5 volts. In general the operating potentials should be low in order that S_g in Eq. (46) may be made a maximum, the limit being determined by the amount of current available to operate the galvanometer.

The use of low operating potentials in addition to increasing sensitivity is of the utmost importance in stabilizing the circuit as ionization of the residual gas in the tube is prevented and leakage currents and drain on the supply batteries are reduced to a minimum. The use of low potentials also allows high capacity storage cells to be employed for all potentials, and consequent to its reduced temperature, alteration of the filament structure with time is reduced to such an extent as to become negligible.

When making practical use of the basic circuit the addition of two slight modifications will be found of value in extending its usefulness and in allowing considerable latitude in the choice of a galvanometer. The modified circuit is reproduced in Fig. 7.

The first addition to the circuit is the application to the galvanometer of a current flowing in the opposite direction to the plate current so that the resultant galvanometer current flow is reduced and a more sensitive instrument may be employed. There are a large number of methods of carrying out this step which will be discussed with different circuits and attention for the moment is only directed to the most obvious which is to employ an auxiliary battery and resistance.

Referring to Fig. 7 it may be shown, after the manner of deriving Eq. (19), that the variation of the galvanometer current with the grid potential is given by the equation

$$\partial i_{ga}e/\partial E_g = K_m/(1+r_{ga}e/r_{cd}), \qquad (48)$$

where r_{cd} is the critical damping resistance which in this case is given by

$$1/r_{cd} = 1/Z_p + 1/r_a + 1/r_d, \qquad (49)$$

r_d being a resistance shunted across the galvanometer to make this instrument critically damped. When low operating potentials are employed with the tube a single dry cell is entirely satisfactory for the compensating battery. r_a may be a ten thousand ohm variable resistance in series with a four hundred ohm variable for fine adjustment. The value of r_d depends upon the galvanometer constants and can be determined from Eq. (49). It is, however, seldom critical and may be conveniently obtained empirically.

The second addition to the circuit is a source of variable known potential applied in the grid circuit in the same direction as the negative bias. This potential is normally kept at zero, and when the input current to be amplified passes across the input resistor causing the generation of a positive potential in the grid circuit, additional negative grid potential is added from the auxiliary source until the galvanometer deflection is restored to its initial position. When this condition is established the e.m.f. from the potentiometer is equal and opposite to that generated by the input current. The absolute value of the input current is then immediately obtainable from the product of the added e.m.f. and the

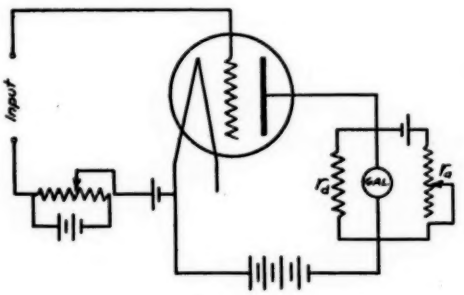

FIG. 7. Simple direct-current amplifying circuit fitted with plate current compensator to allow increased sensitivity of the recording galvanometer. The potentiometer in the grid circuit is used to compensate for the input voltage thus keeping all the potentials of the amplifier constant during the process of amplification.

value of the input resistor. In addition to yielding absolute values of the input current this null method of measurement is of extreme value in that it is not necessary to have the valve characteristics adjusted so that the plate current is directly proportional to the grid potential.

Discussion of the basic circuit has so far been limited to considerations of voltage sensitivity. Maximum sensitivity to current will result when the input resistance, usually limited by the value of the grid impedance, is a maximum together with the other parameters of Eq. (47). In order that this condition may be obtained it becomes important to be able to control the relative positions of the grid potential *vs.* grid current curve and grid potential *vs.* plate current curve with reference to the axis of grid potential. In examining the position of the grid current curve on the grid potential axis, it is convenient and satisfactory to plot the floating grid potential and use this point as indicative of the motion of the whole curve. The actual extent of the variations of tube characteristics depends upon the type of tube employed but the general trend is similar for all tubes and may be gathered from examination of Figs. 8 and 9. The data for these figures have been obtained for the UX–222, a tube that is most satisfactory for many purposes. It has a tungsten filament, is of the screen grid type and the control grid is led out of the apex of the glass envelope and is thus insulated in the best possible manner from the remaining elements. The presence of a screen grid is of use in that it allows some latitude to be exercised in choosing desired characteristics. Fig. 8 deals with plate and screen grid current characteristics of such a tube as functions of plate and screen grid potentials. This information is of value as, in addition to its regular use, the tube may be operated as a triode by connecting the screen grid and plate together, or the screen grid may be employed as a plate, the real plate in this instance being kept at a potential less than that of the screen grid. The variation of the plate current characteristics with the filament potential is reproduced for a constant value of the plate potential in Fig. 9. The locus of floating grid potential is designated in this figure by the broken line. The variation of this potential with the

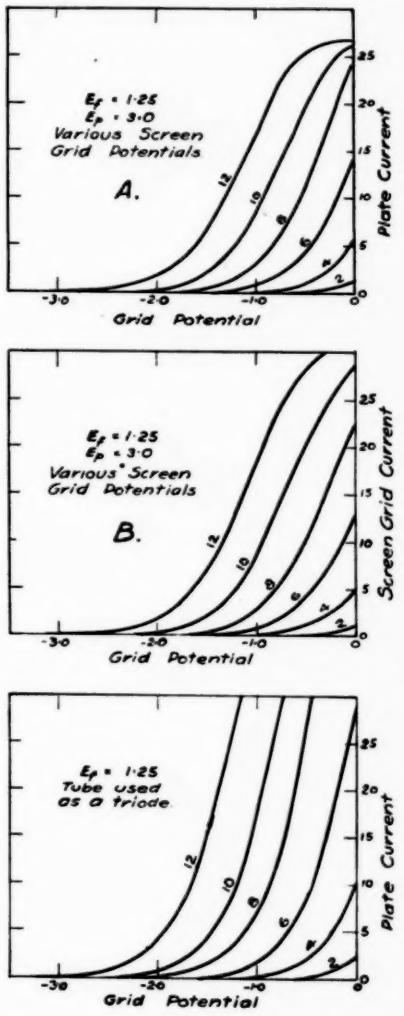

FIG. 8. Characteristics of the UX-222 screen grid tube as a function of the operating potentials. Fig. A shows a study of the plate current, Fig. B the screen grid current and Fig. C the combined screen grid and plate currents, the tube in this instance being used as a triode. The tube may be operated as an efficient amplifier by placing the recording galvanometer in any of these three circuits.

screen grid potential ($E_{sc}>E_p$) is reproduced in Fig. 10. A similar curve is obtained by varying the plate potential when the tube is used as a triode. The general manner of variation of the grid current has already been discussed in the previous section of this paper dealing with the nature of the grid current. This information together with that just given allows a summary of the properties of tube characteristics to be drawn up in so far as they affect the amplification of direct currents.

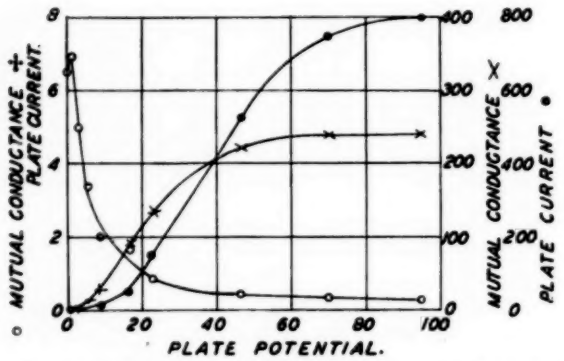

FIG. 6. Data showing, after the manner of Fig. 5, how the sensitivity to voltage increases with decreasing plate potential. The mutual conductance divided by the plate current is proportional to the sensitivity measured in millimeters per volt.

volt/mm with a plate potential of 1.5 volts. In general the operating potentials should be low in order that S_g in Eq. (46) may be made a maximum, the limit being determined by the amount of current available to operate the galvanometer.

The use of low operating potentials in addition to increasing sensitivity is of the utmost importance in stabilizing the circuit as ionization of the residual gas in the tube is prevented and leakage currents and drain on the supply batteries are reduced to a minimum. The use of low potentials also allows high capacity storage cells to be employed for all potentials, and consequent to its reduced temperature, alteration of the filament structure with time is reduced to such an extent as to become negligible.

When making practical use of the basic circuit the addition of two slight modifications will be found of value in extending its usefulness and in allowing considerable latitude in the choice of a galvanometer. The modified circuit is reproduced in Fig. 7.

The first addition to the circuit is the application to the galvanometer of a current flowing in the opposite direction to the plate current so that the resultant galvanometer current flow is reduced and a more sensitive instrument may be employed. There are a large number of methods of carrying out this step which will be discussed with different circuits and attention for the moment is only directed to the most obvious which is to employ an auxiliary battery and resistance.

Referring to Fig. 7 it may be shown, after the manner of deriving Eq. (19), that the variation of the galvanometer current with the grid potential is given by the equation

$$\partial i_{ga}e/\partial E_g = K_m/(1+r_{ga}e/r_{cd}), \qquad (48)$$

where r_{cd} is the critical damping resistance which in this case is given by

$$1/r_{cd} = 1/Z_p + 1/r_a + 1/r_d, \qquad (49)$$

r_d being a resistance shunted across the galvanometer to make this instrument critically damped. When low operating potentials are employed with the tube a single dry cell is entirely satisfactory for the compensating battery. r_a may be a ten thousand ohm variable resistance in series with a four hundred ohm variable for fine adjustment. The value of r_d depends upon the galvanometer constants and can be determined from Eq. (49). It is, however, seldom critical and may be conveniently obtained empirically.

The second addition to the circuit is a source of variable known potential applied in the grid circuit in the same direction as the negative bias. This potential is normally kept at zero, and when the input current to be amplified passes across the input resistor causing the generation of a positive potential in the grid circuit, additional negative grid potential is added from the auxiliary source until the galvanometer deflection is restored to its initial position. When this condition is established the e.m.f. from the potentiometer is equal and opposite to that generated by the input current. The absolute value of the input current is then immediately obtainable from the product of the added e.m.f. and the

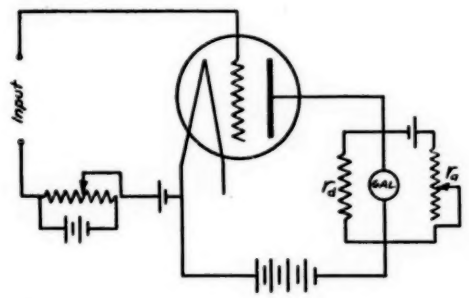

FIG. 7. Simple direct-current amplifying circuit fitted with plate current compensator to allow increased sensitivity of the recording galvanometer. The potentiometer in the grid circuit is used to compensate for the input voltage thus keeping all the potentials of the amplifier constant during the process of amplification.

value of the input resistor. In addition to yielding absolute values of the input current this null method of measurement is of extreme value in that it is not necessary to have the valve characteristics adjusted so that the plate current is directly proportional to the grid potential.

Discussion of the basic circuit has so far been limited to considerations of voltage sensitivity. Maximum sensitivity to current will result when the input resistance, usually limited by the value of the grid impedance, is a maximum together with the other parameters of Eq. (47). In order that this condition may be obtained it becomes important to be able to control the relative positions of the grid potential *vs.* grid current curve and grid potential *vs.* plate current curve with reference to the axis of grid potential. In examining the position of the grid current curve on the grid potential axis, it is convenient and satisfactory to plot the floating grid potential and use this point as indicative of the motion of the whole curve. The actual extent of the variations of tube characteristics depends upon the type of tube employed but the general trend is similar for all tubes and may be gathered from examination of Figs. 8 and 9. The data for these figures have been obtained for the UX–222, a tube that is most satisfactory for many purposes. It has a tungsten filament, is of the screen grid type and the control grid is led out of the apex of the glass envelope and is thus insulated in the best possible manner from the remaining elements. The presence of a screen grid is of use in that it allows some latitude to be exercised in choosing desired characteristics. Fig. 8 deals with plate and screen grid current characteristics of such a tube as functions of plate and screen grid potentials. This information is of value as, in addition to its regular use, the tube may be operated as a triode by connecting the screen grid and plate together, or the screen grid may be employed as a plate, the real plate in this instance being kept at a potential less than that of the screen grid. The variation of the plate current characteristics with the filament potential is reproduced for a constant value of the plate potential in Fig. 9. The locus of floating grid potential is designated in this figure by the broken line. The variation of this potential with the

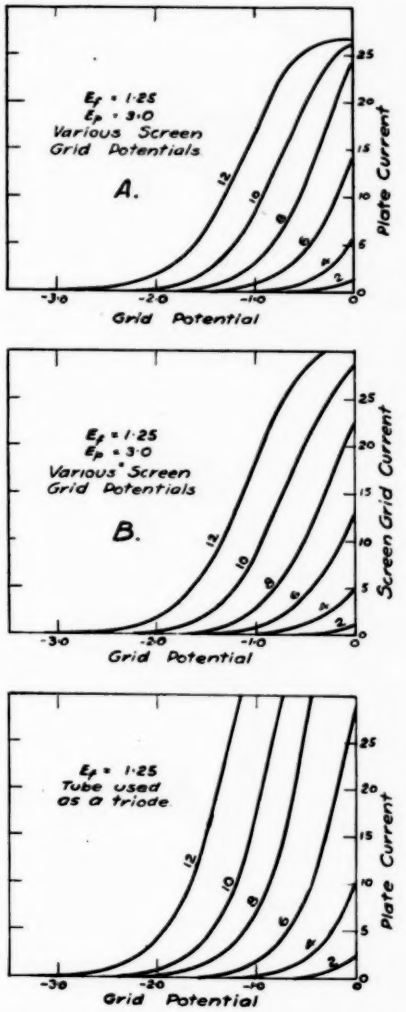

FIG. 8. Characteristics of the UX-222 screen grid tube as a function of the operating potentials. Fig. A shows a study of the plate current, Fig. B the screen grid current and Fig. C the combined screen grid and plate currents, the tube in this instance being used as a triode. The tube may be operated as an efficient amplifier by placing the recording galvanometer in any of these three circuits.

screen grid potential ($E_{sc} > E_p$) is reproduced in Fig. 10. A similar curve is obtained by varying the plate potential when the tube is used as a triode. The general manner of variation of the grid current has already been discussed in the previous section of this paper dealing with the nature of the grid current. This information together with that just given allows a summary of the properties of tube characteristics to be drawn up in so far as they affect the amplification of direct currents.

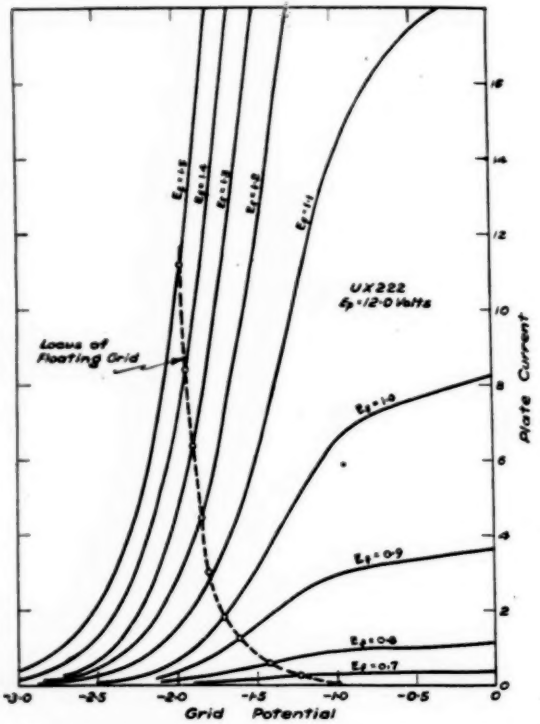

FIG. 9. Characteristics of a UX-222 used as a triode, as a function of the filament potential E_f. It is quite feasible to operate with a filament potential of 0.7 volt in two tube circuits as in this case the low filament voltage in the input tube gives stability and a high grid impedance while the second tube supplies the desired mutual conductance.

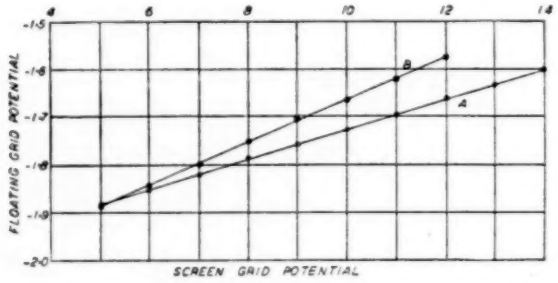

FIG. 10. Data obtained for two different filament voltages which show that the floating grid potential is directly proportional to the screen grid potential. Similar data are obtained for variations of the plate potential with a triode.

SUMMARY OF TUBE CHARACTERISTICS AS FUNCTIONS OF OPERATING POTENTIALS

Reduction of filament potential

Reduces grid current.

Reduces plate current.

Increases stability of circuit by reducing evaporation from filament and drain on supply batteries.

Moves floating grid potential toward axis of zero grid potential.

Moves the grid-potential *vs.* plate-current curve toward axis of zero grid potential.

Provided the plate potential is low the mutual conductance decreases with decreasing filament potential below a specific value; above this value the mutual conductance is independent of the filament potential.

Plate potential

The plate potential should not greatly exceed 12 volts when amplifying small currents owing to the production of gaseous ions and decreased stability at high values. In order to be absolutely free from ionization a plate potential of 8 volts must not be exceeded.

Lowering the plate potential moves the grid-potential *vs.* plate-current curve towards the axis of zero grid potential, the shift being in direct proportion to the decrease in plate potential.

Under many conditions the mutual conductance is practically independent of the plate.

The plate current is decreased with decreasing plate potential.

Floating grid potential varies in direct proportion to the plate potential; the magnitude of the variation is, however, small.

Screen grid potential

With the screen grid potential in excess of the plate potential the potential of this grid affects the tube characteristics in the same manner as the plate potential in a triode.

The grid-potential *vs.* plate-current curve is practically independent of the plate potential provided the latter does not exceed the potential of the screen grid.

As an example of the application of this information attention is directed to the characteristics of the UX-222 tube operated in a screen grid circuit. Since the plate potential is not critical it may be chosen as 1.5 or 2.0 volts depending on whether a dry or storage cell is to be employed for the energy in the plate circuit. The value of the screen grid potential is made as large as possible in order to move the grid-potential *vs.* plate-current curve well into the region of high negative bias, since the grid variational conductance in this region is small. The maximum

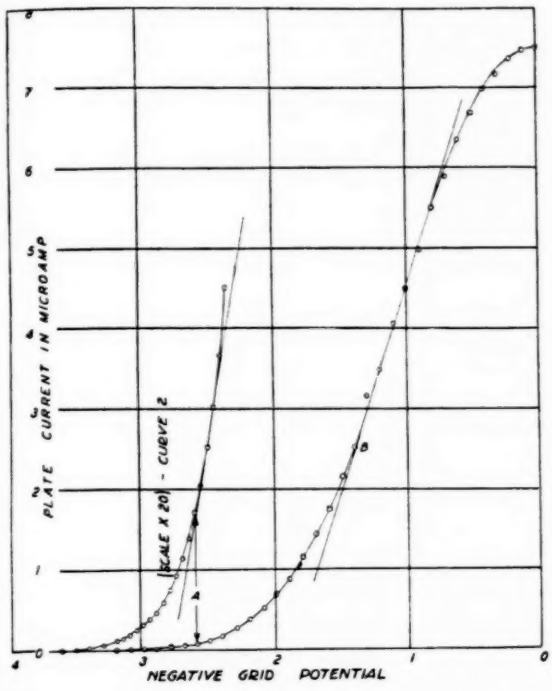

FIG. 11. The plate current characteristic of the simple circuit direct current amplifier having a high current sensitivity. For maximum sensitivity the valve is operated at point A. Curve 2 is the lower section of curve 1 plotted on a larger vertical scale. If a plate current neutralizing circuit is employed the valve is operated at the point B.

value of the screen grid potential consistent with stability is 12 volts, which is accordingly employed. The filament potential is kept as low as is consistent with the graph of Fig. 5 and a measurable plate current. For one particular tube now to be considered in detail this value was chosen as 1.25 volts. The use of this low filament potential ensures stability and freedom from drift. The value of the necessary negative bias depends to a considerable extent upon the characteristics of the specific tube. These are reproduced on a large scale in Fig. 11. The measured quotient of K_m/i_p at the point A is 5.0, at B it is 2.3; accordingly if there is no plate current compensator across the galvanometer, greater voltage sensitivity will result by operating at A. The grid impedance at A is 10^{14} ohms with a positive ion grid current, while at B the impedance is 10^{11} ohms and is just at floating grid potential. Curve 2 of this graph shows that K_m is almost linear for approximately 0.2 volt above point A. The operating constants of the tube operated at the point A are given in Table I, column 1.

TABLE I

Column 1 gives the operating constants of a UX-222 used in the simple circuit of Fig. 2B. Column 2 gives the operating constants of the same tube used as a floating grid amplifier according to Fig. 12.

	Column 1	Column 2
Filament Voltage	1.25 volts	1.20 volts
Plate Voltage	1.5 volts	3.00 volts
Galvanometer Sensitivity	3×10^{-10} amp./mm	10^{-10} amp./mm
Voltage Sensitivity	8×10^{-4} volt/mm	10^{-5} volt/div.
Plate Current	0.07×10^{-6} amp.	——
Mutual Conductance	0.35×10^{-6} amp./volt	——
Screen Grid Potential	12.0 volts	12.5 volts
Grid Impedance	10^{14} ohms (positive ion)	10^{11} ohms

The grid impedance in this instance is greater than the value necessary to utilize the maximum input resistor of 10^{11} ohms. It is, as has already been explained, desirable to have this grid impedance large relative to the input resistor so that any departure from linearity in the former will result in a minimum departure from the linear response of the galvanometer deflection in relation to the input current. Where high precision is required it is desirable to use a grid potential compensating potentiometer and a null method of measurement.

The amplifier may be operated at floating grid potential employing the circuit of Fig. 12. Since,

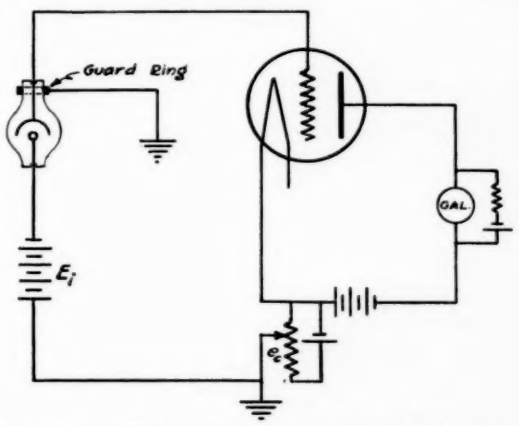

FIG. 12. Amplifier operated at floating grid potential for photoelectric currents. The guard ring is connected across the photoelectric cell and supply battery to prevent leakage currents flowing to the grid. This circuit may be operated as a null instrument by setting the potential of E_i above the saturation potential of the photocell and compensating for the photoelectric current by continual variation of the potentiometer e_c to keep the galvanometer deflection constant.

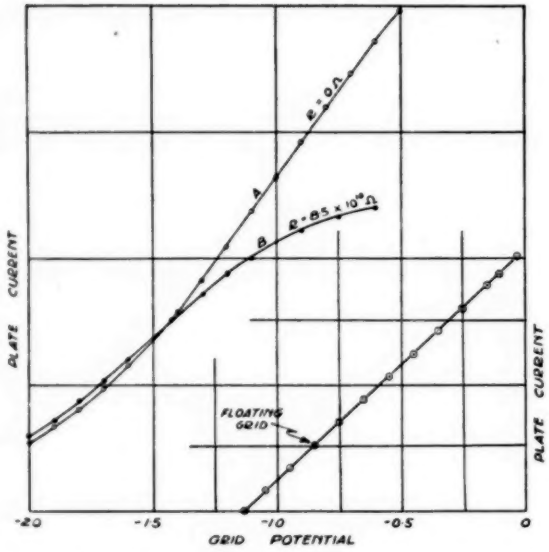

FIG. 13. Data showing the extent of linear grid impedance in the vicinity of the floating grid potential. The grid impedance is constant over the grid potential range in which curves A and B are both straight. Curve C is a large scale plotting of B near floating grid potential.

however, a null method cannot conveniently be employed with this circuit it is essential that both K_m and K_g be linear over the working range employed. The graphs of Fig. 13 demonstrate the simplest method of determining the extent of the linear section of K_g by employing Eq. (33). Referring to this figure, curve A is the grid-potential *vs.* plate-current curve run with no resistor in the grid circuit. Curve B is a repetition of A with a resistor of 8.5×10^{10} ohms included in series in the grid circuit. According to Eq. (33) K_g is constant as long as the two curves are straight lines. From the graph this range measures in excess of 0.2 volt to the right of the floating grid potential. The exactitude of the linear relation is clearly shown by curve C which is an enlarged plotting of B near the floating grid region. According to these data the amplifier response will bear a linear relation to the input current as long as this current does not raise the potential of the grid in excess of 0.2 volt above floating grid potential.

When employing a floating grid circuit of this type to measure photoelectric currents care must be taken to see that there is no leakage or dark current across the cell, as the presence of such a current will force the grid potential away from floating grid and even off the section of linear response. Dark currents arise from thermionic emission from the photoelectric surface of some types of photocells and are usually overcome by reducing the temperature of the cell. Fortunately, however, there are many types of cells free from this effect. Leakage currents are due to charges flowing over imperfectly insulating surfaces and often reach an appreciable size because of the high voltage it is necessary to apply across a photoelectric cell. Such currents may be eliminated by the use of guard rings, which are contacts made around the photocell both within and outside and which are kept at the same potential as the grid thus preventing any stray current flow to this element. The manner of connecting a guard ring is shown in Fig. 12. In this figure there is also shown a potentiometer connected in a manner equivalent to a guard ring for the grid lead. This potentiometer allows the grid to be operated at zero potential relative to the ground by raising the potential of the negative side of the filament an amount equal to the negative bias. When this condition has been established it is possible to employ a long lead between the photocell and control grid, since the leakage from this lead will then have been reduced to a minimum. Floating grid circuits almost always require the use of a compensator across the galvanometer since the necessity of operating on the region of linear mutual conductance is accompanied by an appreciable initial plate current. Characteristic data for a UX–222 used as a floating grid amplifier are given in Table I, column 2. The particular instrument from which these data were obtained gave absolute linearity of response for photoelectric currents between 10^{-15} and 10^{-12} amp. An ordinary dry cell was used for compensation in the plate circuit and a motor car storage battery supplied the remainder of the energy.

Space charge grid tubes

Considerable space has been given to discussing the properties of screen grid and triode tubes. Variations of the constants of a space charge tube with applied potentials are in many respects similar. Screen grid tubes may, of course, be used as space charge grid tubes by applying a positive potential to the inner grid and employing the former screen grid as a control grid. A tube em-

ployed in this manner is not suitable for measuring extremely small currents owing to the manner in which the leads are brought out of the glass envelope.

At least two different manufacturers supply space charge grid tubes with the control grid highly insulated and developed especially for the measurement of extremely small currents. The author has had experience only with the FP-54 which is the tube manufactured by the General Electric Company and attention is limited to the discussion of its characteristics. This discussion is probably immediately applicable to another tube manufactured by the Western Electric Company.

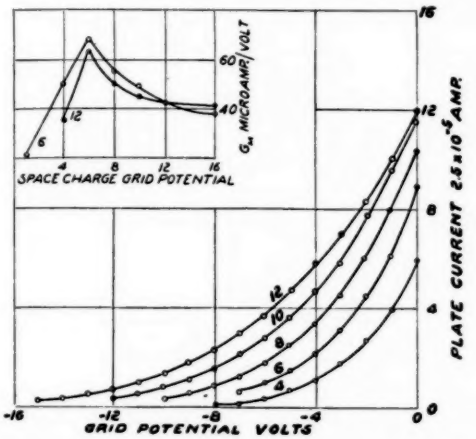

FIG. 14. Data for the FP-54 space charge grid tube showing the dependence of the mutual conductance upon a space charge grid potential. The numerals upon the family of plate current curves denote the space charge grid potential.

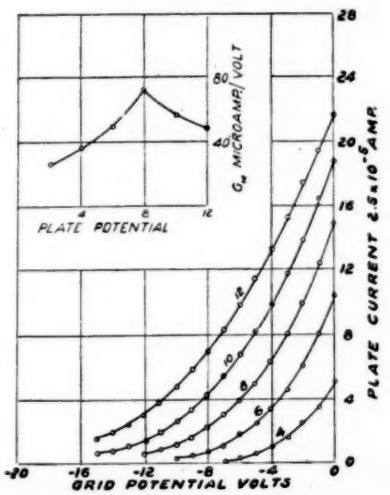

FIG. 15. Data for the FP-54 space charge grid tube showing the manner in which the mutual conductance varies with the plate potential.

The effect of the space charge grid potential upon the mutual conductance is shown in Fig. 14: the effect of the plate potential in Fig. 15. Examination of the grid impedance as a function of the space charge grid potential shows that the grid impedance decreases with increase of this potential but that for all practical voltages it remains in excess of 10^{14} ohms and is thus safely above the value required for satisfactory operation with input resistors up to 10^{11} ohms in value. The foregoing figures were obtained with a filament potential of 2.5 volts which is the maximum allowable on this tube. Reduction of this potential has the same effect upon all characteristics as it has with a triode tube. Operating potentials for the tube used in direct deflection circuits are given in Table II, column 1. In addition to the ready control of the grid current a space grid tube is of extreme value due to the relative position of the space grid and plate with respect to the control grid. The space grid, lying as it does between the control grid and filament, causes a decrease in the current through the space grid circuit with increase of control grid potential, while the same

TABLE II

Column 1 gives the operating constants of an FP-54 direct current amplifying tube that are most satisfactory in direct deflection circuits. Column 2 gives the constants of a vacuum tube electrometer constructed with an FP-54 and employing the circuit of Fig. 2C with the addition of the plate current compensator of Fig. 7.

	Column 1	Column 2
Filament potential	2.5 volts	1.2 volts
Plate potential	8.0 volts	6.0 volts
Space Charge Grid Potential	6.0 volts	3.0 volts
Initial Negative Bias of Control Grid	——	6.0 volts
Operating Grid Current	——	2×10^{-17} amp.
Grid Impedance at Floating Grid Potential	——	10^{16} ohms
Operating Grid Impedance	10^{14} ohms	10^{18} ohms
Range of Operating Grid Impedance	——	1.0 volt
Voltage Sensitivity	——	2×10^{-3} volt/mm
Mutual Conductance	70×10^{-6} amp./volt	5×10^{-8} amp./volt
Galvanometer Sensitivity	——	10^{-10} amp./mm
Time required for Grid Current to Charge Grid over Range of Operating Grid Impedance	——	10^{5} seconds
Capacity of Condenser System	——	3×10^{-12} farad

increase in this potential causes an increase in the plate current. This phenomenon will later be shown to be the basis of an extremely stable method of obtaining a balanced plate current circuit.

Battery neutralizing circuit

Fluctuations in circuits resulting from instability in supply potentials may be eliminated by the introduction of suitable resistances in various parts of the circuit and the subject has been summarized by Turner.[16] Fig. 16A shows a method suggested by Dearle and Matheson[17] for compensation of filament supply fluctuations. Fig. 16B shows the same principle applied to the plate supply and Fig. 16C shows a completely neutralized circuit.

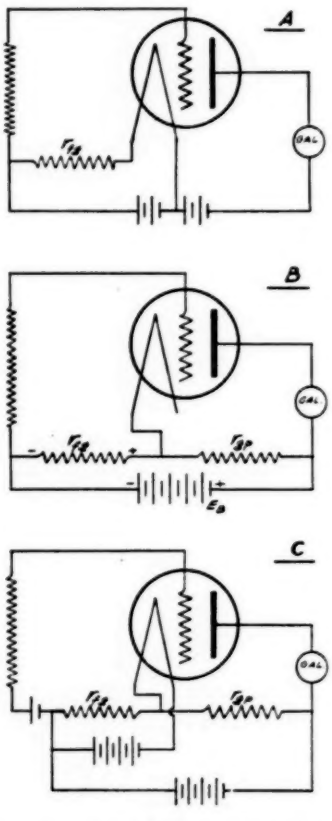

FIG. 16. Battery neutralizing circuit. The resistor r_{fg} in Fig. A compensates for fluctuations in the filament supply battery. The two resistors in Fig. B compensate for fluctuations in the plate supply battery. Fig. C shows the completely neutralized circuit which is extremely stable.

[16] L. A. Turner, Rev. Sci. Inst. **4**, 665 (1933).
[17] R. C. Dearle and L. A. Matheson, Rev. Sci. Inst. **1**, 215 (1930).

The condition for filament supply neutralization is obtained from the equation

$$\Delta i_p = (\partial i_p/\partial i_f)\Delta i_f - K_m r_{fg}\Delta i_f,$$

which is zero when

$$r_{fg} = (\partial i_p/\partial i_f)(1/K_m). \tag{50}$$

Plate supply neutralization is secured when

$$r_{gp} = \mu r_{fg},$$

for under this condition a change ΔE in the plate supply voltage alters the plate current by an amount

$$\Delta i_p = (\mu/(\mu+1))\Delta E$$

and the grid potential by a corresponding amount

$$\Delta E_g = (1/(\mu+1))\Delta E, \tag{51}$$

so that the resulting change in plate current is zero.

Complete neutralization is accordingly obtained by adjusting r_{fg} for the filament condition and then setting r_{gp} according to Eq. (51).

Bridge circuits

The principles involved in the common Wheatstone bridge have been made the basis of most effective balanced plate current circuits of value both for neutralizing fluctuations in supply batteries and for eliminating the initial plate current in the galvanometer circuit. In a simple Wheatstone bridge the condition required for no current

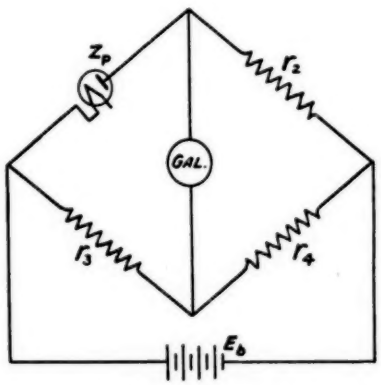

FIG. 17. Diagram showing the filanent to plate section of a valve incorporated in a Wheatstone bridge. When the bridge is balanced all the plate current passes through the resistor r_2. A change in grid potential causes approximately one half the change in galvanometer deflection that is obtained in the simple circuit. This loss in sensitivity is more than offset by the increase in galvanometer sensitivity that may be utilized in the bridge circuit.

to flow through the galvanometer is $r_1r_4=r_2r_3$. Since this equation contains no reference to potential or current the condition for zero current in the galvanometer is independent of these factors. Fig. 17 indicates the manner in which the plate circuit of a tube may be incorporated into the arm of such a bridge. In this circuit the galvanometer zero will be independent of fluctuations in the plate battery when the plate impedance obeys Ohm's law, a condition that can be brought about by operating on the plate-potential *vs.* plate-current curve at the point where the extension of the tangent passes through the origin. This operating potential has disadvantages that are immediately apparent and a number of other conditions for independence of fluctuations of the plate potential may be established.

The current through the galvanometer may be deduced analytically in the same manner as the corresponding calculation for a simple Wheatstone bridge by utilizing the expression for the plate impedance defined by Eq. (16) in place of a pure resistance expression. The resulting value is

$$i_{ga}e=\frac{(r_2r_3+r_2r_4)(E_0+\mu E_g)-(r_4Z_p-r_2r_3)E_B}{r_2Z_pr_3+r_4Z_pr_3+r_2r_4Z_p+r_2r_4r_3+r_{ga}e(Z_p+r_2)(r_3+r_4)}. \quad (52)$$

If Z_p, μ, e_0 are independent of E_B then

$$\partial i_{ga}e/\partial E_B=r_4Z_p-r_2r_3. \quad (53)$$

For this to be zero

$$r_4Z_p=r_2r_3. \quad (54)$$

If now it be assumed that $r_2\doteq r_4$ and $Z_p\doteq r_3$, a condition employed in practice, then

$$\partial i_{ga}e/\partial E_g=K_m/[2+r_{ga}e(1/Z_p+1/r_2)] \quad (55)$$

or, if as in the circuit of Fig. 7, a resistance r_d be shunted across the galvanometer to damp this instrument critically, the equation may be written:

$$\partial i_{ga}e/\partial E_g=K_m/2(1+r_{ga}e/r_{cd}), \quad (56)$$

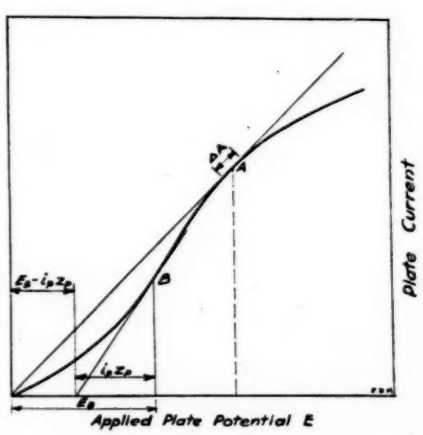

FIG. 18. The plate current characteristic of a valve. Operation at the point A results in the plate impedance having the same characteristics as an ohmic resistance. At the point B the plate impedance is equivalent to an ohmic resistance in series with an e.m.f. of value $E_B-i_pZ_p$.

which equation shows that the sensitivity of a bridge circuit is one-half that of the simple type, this result arising from the existence of two alternative current paths about the recording galvanometer.

Conditions of independence of the bridge circuit for variations of E_B have been summarized by a number of investigators.[11, 17, 19] It is sufficient for the purposes of this communication to discuss only the principles involved. Referring to Fig. 18 showing the relation between the plate potential and plate current, it will be seen that should the potential at A be chosen for operation, then since the tangent at this point passes through the origin, the tube will behave as a pure resistance over the limited range ΔA. Operating at the point B is equivalent to placing in a bridge arm a resistance of value Z_p in series with an e.m.f. of value $E_B-i_pZ_p$. The behavior of the bridge in this case may be deduced analytically and shown to be independent of plate supply fluctuations over a definite range but the resulting circuit involves the addition of an external battery which destroys its simplicity. Fluctuations are often reduced by employing two tubes in a balanced bridge circuit after the manner of Fig. 19. In this circuit originally devised by Wold[18] and developed by Wynn-Williams,[19] emission from the two filaments is adjusted by the rheostat r_f until a change in the filament

[18] P. I. Wold, U. S. Patent, 1,232,879 (1917).

[19] C. E. Wynn-Williams, Proc. Camb. Phil. Soc. **23**, 810 (1927).

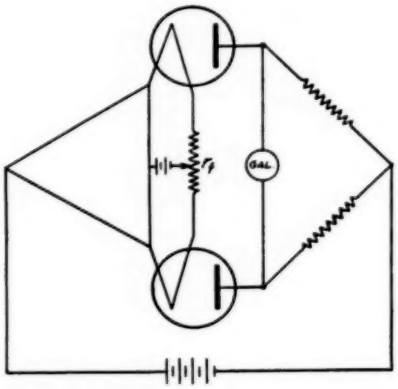

FIG. 19. Two tube balanced bridge circuit. The resistor r_f allows the filament emission characteristics of the two tubes to be made identical. Conditions for exact plate supply compensation have been developed by a number of investigators. Prior to the development of the space charge balanced bridge circuit this type of circuit represented the most sensitive and stable amplifier.

supply causes a minimum alteration in the galvanometer current; that is to say, the filament-current *vs.* plate-current characteristics of the two tubes are so matched that over a small range a change in filament supply affects both tubes in the same manner.

While the foregoing circuits are of some theoretical interest the most satisfactory type of bridge circuit is one in which a common battery is used to supply all power, the filament circuit actually constituting one section of the potential divider. A circuit of this type was published by Soller[20] and has since been developed in many forms.

Precision single tube space charge grid balanced circuit

DuBridge and Brown[21] incorporating Soller's principle have described a circuit in which the space charge grid and plate circuits of the same tube are made opposing arms of the same Wheatstone bridge. This beautifully simple adaptation of the bridge system yields an ideal circuit, as compensation for fluctuations in filament emission arising not only from variation in the supply potentials, but also from actual changes in the physical characteristics of the filament, is provided. It is at present the only circuit capable of this most important form of compensation and as such represents a basic improvement on all others as alterations in the physical structure of the filament have to be recognized as one of the basic weaknesses of a thermionic circuit.

The removal of the effects of these filament changes enables a circuit to be readily operated at its theoretical limit. Fig. 20A reproduces the circuit. Referring to this figure it will be seen that the drop in potential across the galvanometer is

$$\Delta E_{ga}e = r_s i_s - r_p i_p. \tag{57}$$

When this potential is zero

$$r_s/r_p = i_p/i_s. \tag{58}$$

Any fluctuation in the supply battery will cause a change in the filament current so that independdence of such fluctuation is obtained when

$$\partial(\Delta E_{ga}e)/\partial i_f = 0. \tag{59}$$

Incorporating this condition in (57) yields

$$\partial i_p/\partial i_f = (r_s/r_p)(\partial i_s/\partial i_f). \tag{60}$$

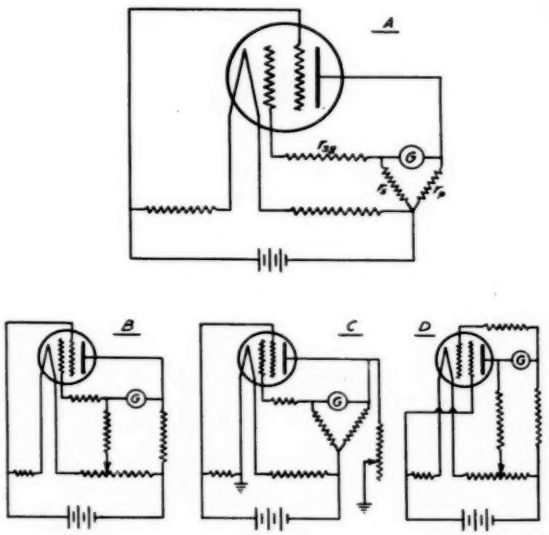

FIG. 20. Single tube space charge balanced bridge circuit. The current passing through the space charge grid circuit neutralizes the effect upon the galvanometer of the normal plate current. A rise in control grid potential increases the plate current and decreases the space charge grid current. Fig. A is the original circuit of DuBridge and Brown in which both space currents are derived from the same filament thus enabling the effect of changes in filament emission to be eliminated. The lower three circuits are modifications by different investigators of the same circuit and developed to enable the space grid and plate circuit characteristics to be readily matched. These modified circuits all allow some current, which has not been transmitted from the filament, to go through the galvanometer.

[20] W. Soller, Rev. Sci. Inst. **3**, 416 (1932).

[21] L. A. DuBridge and H. Brown, Rev. Sci. Inst. **4**, 532 (1933).

In order that (58) and (60) be satisfied it is necessary that the i_p *vs.* i_f and the i_s *vs.* i_f curves be linear over the working range and that they intersect on the i_f axis. To attain such conditions in practice the operating constants of the tube are set with approximately the desired filament current, the two curves in question determined and the correct filament current obtained. The function of the resistance r_{sg} is to enable the correct potential to be applied between the filament and space charge grid.

A number of investigators have examined this circuit in attempts to make the conditions of balance more flexible. Harnwell and Van Voorhis[22] have developed a mathematical analysis expressing the conditions of balance in terms of the circuit parameters and this has led to the introduction of auxiliary resistances in various parts of the circuit. The introduction of these resistances is in all cases equivalent to introducing somewhere in the balanced circuit a source of e.m.f. which is not dependent upon variation of the filament structure and for this reason the present author does not feel that they represent a true advance in the development of the circuit. An excellent summary of these various circuits is given by Penick[23] who includes a modification of his own. The circuits as they have been discussed are reproduced in Fig. 20, from a study of which the functions of the various resistors in yielding flexibility for adapting the circuit to different tube constants will be at once apparent.

The vacuum tube electrometer

An electrometer is an instrument for measuring current by observing the rate of change of potential across a condenser resulting from an accumulation of charge derived from the input current. The basic equation governing such an instrument is accordingly

$$dE/dt=(1/C)(dQ/dT)=i/C, \qquad (61)$$

where i is the input current. The common forms of this instrument are mechanical in their action in so far as they register the potential change by the motion of a part of the condenser flexibly mounted in such a manner as to readily react to a changing electrostatic force.

When a vacuum tube is employed as an electrometer the grid-filament system constitutes the condenser (Fig. 2C) and the rate of accumulation of its charge is obtained by observing the rate of change of the plate current. The basic equation is accordingly

$$dE_g/dt=(1/C)(i\pm i_g). \qquad (62)$$

The positive sign is usually employed in front of i_g as it is almost always necessary to operate on the positive ion section of the grid current curve, which current will charge the condenser in the same direction as the input current.

In order that a vacuum tube may have the same electrical properties as a mechanical electrometer it is necessary that i_g in Eq. (62) be made zero. This condition has not yet been completely attained but it is readily possible to make the quantity so small that for nearly all purposes it may either be ignored or an exact correction added.

In order to reduce i_g to a minimum it is only necessary to reduce the filament potential. A family of characteristics of the FP-54 suitable for electrometer use is reproduced in Fig. 21. The intermittent line plotted with crosses in the graph represents a grid current curve obtained with a plate potential of 6 volts and a space charge grid potential of 3.0 volts. Practically an identical curve is obtained with the plate at 6 volts and a space charge grid potential of 2 volts, so that this latter value is not at all critical. Plate potentials less than 6 volts reduce the plate current to an undesirable extent and do not appreciably alter the grid current characteristics so that it may be concluded that the data of Fig. 21 give the most desirable characteristics of the tube when used as an electrometer.

The voltage sensitivity of this system will be governed by the sensitivity of the galvanometer in the plate circuit which sensitivity may be controlled by adopting one of the plate current sections of a direct deflection circuit.

The properties of a single tube electrometer constructed on the basis of the foregoing data and employing the plate circuit of Fig. 7 are given in Table II, column 2.

The expression for sensitivity of an electrometer differs from that of a direct deflection instrument in that the time required for a read-

[22] G. P. Harnwell and S. N. Van Voorhis, Rev. Sci. Inst. **5**, 244 (1934).
[23] D. B. Penick, Rev. Sci. Inst. **6**, 115 (1935).

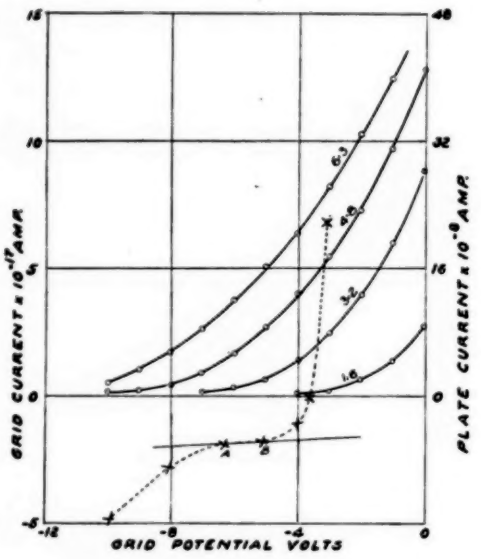

FIG. 21. Family of tube characteristics of the FP-54 used as an electrometer. The circled lines are grid potential *vs.* plate current curves for different plate potentials. The grid current curve is designated by crosses. The grid current has been reduced to a minimum by reducing the filament temperature. The resulting reduction in plate current is no great disadvantage as it is still sufficiently large to operate a standard galvanometer.

ing is involved, is arbitrary and must accordingly be defined. It is convenient to express the current sensitivity as the amount of current required to cause the plate current meter to change one division in sixty seconds. The above amplifier records a current of 10^{-16} amp. in sixty seconds. The use of an improved telescope and scale with the same galvanometer increases the sensitivity to 2×10^{-17} amp. without difficulty in assembling and with absolutely no loss in stability. When currents of this extremely small value are being measured allowance must be made for the grid current. Such a correction is readily applied with a high degree of accuracy if the tube is operated between the points A and B on the grid current curve (Fig. 21), for over this range the grid current is constant. Under these circumstances it is also somewhat better practice to cause the input current to lower the grid potential thus counteracting the grid current charge. The exact extent to which the grid potential has changed is obtained by noting the initial and final positions of the galvanometer, closing the key K and moving the galvanometer over the same range by manipulation of the potentiometer e_c (Fig. 2C).

Elaborate two tube bridge circuits have been described for electrometer use in which both grid circuits are opened at the same instant and the galvanometer deflection resulting from the grid currents balanced out by introducing a small variable condenser across the balancing tube and adjusting its capacity until both grids charge up at identical rates. This procedure is undesirable and may be misleading since the balancing does not remove the grid currents but only masks their effect. Any error from a change in the value of the grid current will still be introduced and due to the presence of two grid currents the chances of this error are doubled. A second and very marked disadvantage of this type of circuit is that the operator has no means of ascertaining when the grid has passed off the range of linear grid current. Single tube circuits then are to be preferred and the most stable of these circuits is that of DuBridge and Brown, though the auxiliary battery circuit of Fig. 7 is suitable for many purposes.

A null method of measurement may be employed by using the potentiometer e_c (Fig. 2C) to alter continuously the potential across the grid condenser system at such a rate as to keep the galvanometer deflection constant. Under this condition the charge upon the grid condenser resulting from the input current will, according to Eq. (61), be exactly nullified by the loss of charge resulting from the reduction of the potential by the potentiometer e_c. After a period T the absolute value of the input current will be EC_g/T where E is the potential registered by the potentiometer e_c. When very small currents are being measured the addition of a special condenser is not desirable but the same result may be obtained by employing the capacity inherent in the photoelectric cell or ionization chamber in which the current originates. To adopt this method it is only necessary to set the potential of the battery E_i above the saturation potential of the ionization cell and insert the potentiometer e_c in series with this battery. Such an arrangement gives an extremely sensitive null circuit.

With regard to the relative sensitivity and ease of operation of the mechanical and vacuum tube electrometers it would seem that the vacuum tube electrometer is considerably superior. The capacity of the tube electrometer is small, the

batteries and galvanometer may be mounted at a considerable distance from the tube and the mechanical parts of a galvanometer are a great deal more rugged than those of an electrometer. The grid current which at first sight seems a great handicap offers in reality very little trouble as it is so readily reduced and remains constant while the electrometer has an equivalent defect due to physical changes in its delicate suspension. The design by Metcalf and Thomson and release of the FP-54 electrometer tube by the General Electric Company marked the point at which the vacuum tube electrometer rose from an object of interest to an instrument of precision.

Negative impedance circuit

The relation between grid potential and plate current in a circuit containing a plate resistance r_p has been shown to be

$$\partial i_p/\partial E_g = K_m/(1+r_p/Z_p). \qquad (19)$$

If Z_p were negative and approached r_p in numerical value the numerical value of this expression would approach infinity.

The existence of negative values for the grid impedance has been pointed out. If then the grid circuit of one tube be connected in series in the plate circuit of another it should be possible to obtain a high effective mutual conductance by making the two impedances approach each other in numerical value and keeping the algebraic signs opposite. Such a circuit has been discussed by Macdonald and Macpherson[24] and Macdonald and Campbell.[25] The circuit in its simplest form is reproduced in Fig. 22 where numerical equality of the two impedances is brought about by adjusting the operating potentials of the two tubes; the battery E in this circuit serves the dual purpose of supplying the grid potential of the output tube and the plate potential of the input tube. In general the positive plate impedance of the tube T_1 is numerically much less than the negative grid impedance of the tube T_2 so that equality is brought about by using low potentials on T_1 to increase this impedance and high potentials on T_2 to decrease the grid impedance. Experiment shows that, with the latter tube, positive ion emission from the filament is not sufficient to enable a condition of equality to be approached and the positive ion current must be increased by increasing the plate potential to a point where gaseous ions are formed. In the interests of stability this value must be kept as low as possible and an applied plate potential of 22.5 volts will generally be found sufficient to ensure satisfactory results.

In considering this circuit it must be realized that the grid of T_2 may never have a potential more positive than the floating grid potential of this tube for if the grid impedance becomes positive all the potential of the battery E will be impressed across T_1 due to the high resistance at this point and the grid of T_2 will assume the floating grid potential resulting from the battery A_2. Further in order for the circuit to operate as an amplifier the grid impedance of the output tube must be negative and hence the potential of this grid must be sufficiently below floating grid to operate upon the negative portion of the characteristic. This limitation upon the characteristic of T_2 at once limits the choice of a suitable output tube to one in which there is an appreciable mutual conductance on the positive ion side of the grid potential. In order for the tube T_1 to have a high plate impedance a tube of the screen grid type is best employed. The use of a screen grid in this tube is also of value in so far as the linear region of mutual conductance of such a tube is moved to regions of greater negative bias with increasing potential of this grid.

A combination of tubes that at first sight might seem ideal is a low grid current tube such as the

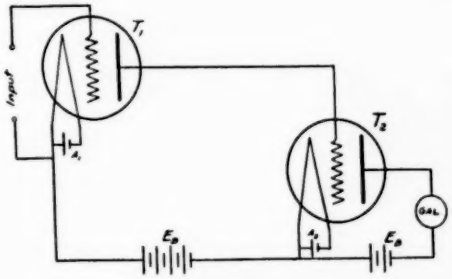

FIG. 22. Amplifier employing the negative impedance characteristic of a grid circuit to obtain a high effective mutual conductance. The negative impedance of the grid circuit of T_2 is connected in series with the positive impedance of the plate circuit of T_1 and the numerical equality of the impedances approached by adjusting the values of E_B and the filament voltage of the tube T_2.

[24] P. A. Macdonald and J. T. Macpherson, Phil. Mag. [7], **15**, 72 (1933).
[25] P. A. Macdonald and E. M. Campbell, Physics **4**, 237 (1933).

FP-54 in the input stage and a tube with a high grid current in the output stage. Experiment shows that a GE-230 connected in the plate circuit of an FP-54 allows a balance of impedances and a set of characteristics are reproduced in Fig. 23. With reference to this figure it is to be noted that the filament potential of the input tube is only 0.68 volt so that the input grid impedance must be very great indeed. It is not however found practicable to employ such a combination as an electrometer and utilize the full value of this impedance since the drift in the plate current of the 230 is appreciable. While then this electrometer is not of as great practical value as others previously discussed, it is, on paper, an ideal instrument and its practical development awaits the production of a suitable two tube combination devoid of drift. The sensitivity limit is accordingly placed at present by practical and not theoretical considerations.

A fact of considerable interest is brought out from a study of the characteristics obtained with a UX-222 input tube and a 230 in the output circuit, for this combination is found to work successfully with a filament potential of 0.6 volt

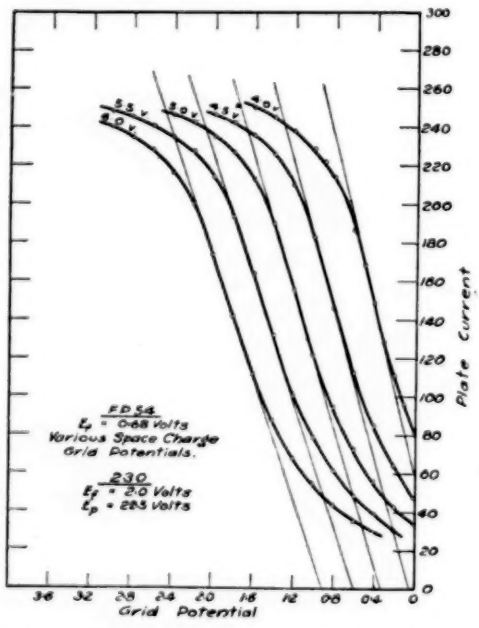

FIG. 23. Characteristics for a negative impedance circuit employing an FP-54 input tube and a 230 in the output stage. The low filament potential of the input tube yields an extremely high input impedance but the drift inherent in the 230 prevents the unit being employed to yield a high sensitivity electrometer.

on the 222. The filament of this tube is not thoriated so that the results suggest a possible method of obtaining a two tube electrometer in which thorium contamination from the filament has been eliminated.

Turning from theoretical to practical considerations, an excellent direct deflection circuit of general utility may be obtained with a 222 and a 230. Such a circuit may be operated entirely from dry cells, including the filament supply.

In addition to use in the laboratory this instrument makes an excellent portable amplifier of high current sensitivity. For a more detailed discussion of this type of circuit and certain tube combinations that may be employed the reader is referred to the original publications.

Two tube circuits

When it is not desired to employ a sensitive galvanometer in the plate circuit, the output current may be forced across a resistor and the resulting voltage drop impressed upon the grid system of a second tube operated with sufficiently high potentials to enable a microammeter to be employed as a plate current indicator. For direct deflection circuits such two tube combinations are excellent, though they are not satisfactory for sensitive electrometer work as there is almost invariably a slow drift of the plate current in the second tube due to the higher operating potentials employed. The key to successful construction of two tube circuits lies in the intelligent selection of the coupling resistor r_p, Fig. 24. This resistor must be made as large as is consistent with the limit placed by the drop in potential occurring across it as a result of the passage of the initial plate current. It is not practicable to allow this drop to exceed the drop across the plate impedance Z_{p_1} of the input tube, as beyond this value the supply potentials required become so great that instabilities are introduced which more than offset any apparent advantage gained. The manner of selecting the coupling resistor is clearly brought out by examination of the data of Fig. 25 showing the characteristics of an FP-54 coupled to a 230 as an output tube. Fig. 25A shows data obtained with an input filament potential of 2.5 volts while Fig. 25B shows data from the same circuit with the input filament potential reduced to 1.5 volts. The operating

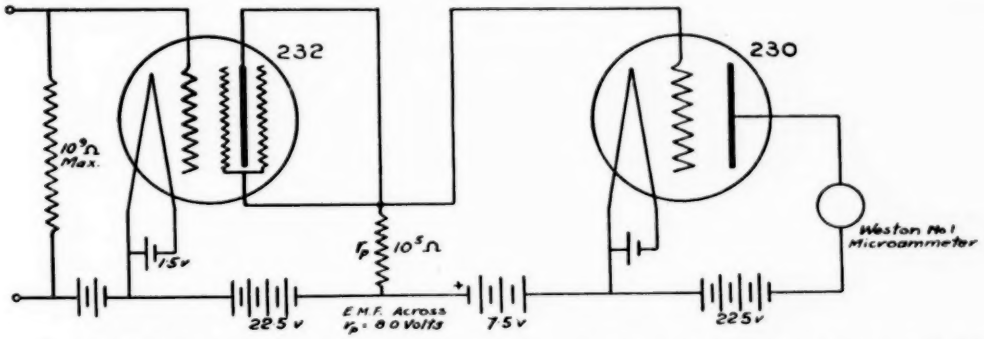

FIG. 24. Two tube resistance coupled circuit. The essential factor for success with such a unit is the choice of r_p. This resistance should be of the maximum value permitted by the voltage drop across it due to the normal plate current. A positive bias battery must be introduced into the grid circuit of the output tube to offset the negative e.m.f. generated across r_p. Reduction of the mutual conductance of the input tube, due to lowering the filament potential, may be offset by increasing the value of the coupling resistor. The numerical values given in this figure yield an excellent low sensitivity two tube circuit (10^{-12} amp./div.).

data are reproduced in Table III from which it will be apparent that a coupling resistor of 10^4 ohms should be used in the first instance and 10^6 ohms in the second. These data clearly show that there is no loss of efficiency in operating the input tube with a low filament potential, in fact this particular circuit operated with complete satisfaction with the input filament at 1.0 volt and a coupling resistor of 5×10^7 ohms.

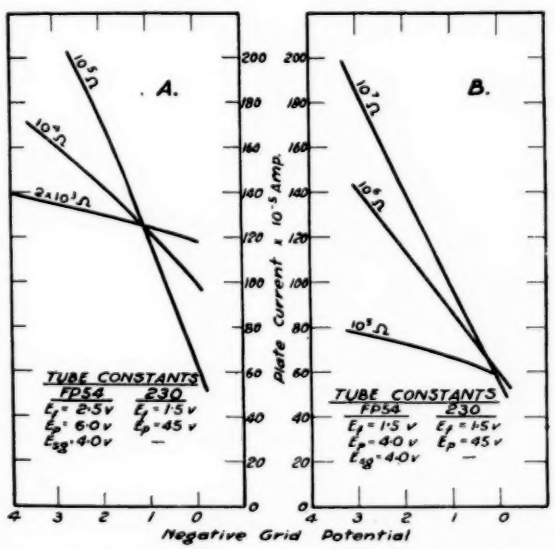

FIG. 25. Operating characteristics of a two tube resistance coupled high sensitivity amplifier. The curves on the left are obtained with a filament potential of 2.5 volts, on the right with 1.5 volts. The numerical values on the curves give the value of the coupling resistor and it will be seen that the increase in the effective mutual conductance of the unit is unaffected by lowering the filament potential, since the coupling resistor may be increased at lower values.

TABLE III. *Operating data for an FP-54 coupled to a 230 as an output tube.*

Filament Potential	Applied Plate Potential	Plate Potential	Coupling Resistor
E_f	E_B	E_p	r_p
2.5	6.5	6.0	2×10^3 ohms
2.5	7.5	6.0	10^4 ohms
2.5	15.0	6.0	10^5 ohms
2.5	68.0	6.0	10^6 ohms
1.5	7.0	6.0	10^5 ohms
1.5	8.0	6.0	10^6 ohms
1.5	20.0	6.0	10^7 omhs

The operating characteristics of the output tube are not critical since the grid impedance of this tube is always considerably in excess of the coupling resistor r_p. It is desirable to keep the plate potential as low as possible and 22.5 volts will almost invariably be found satisfactory. For some circuits it is possible to adjust the applied plate potential of the input tube in such a manner that the drop across the coupling resistor gives the correct operating bias for the output tube. In many instances however, this value exceeds the correct value and it becomes necessary to supply an opposing positive e.m.f. in the grid circuit of this tube. Fig. 24 shows an excellent two tube circuit with a current sensitivity of 10^{-12} amp./mm and which clearly shows how the various potentials of the two tubes are adjusted. Two tube circuits are not employed nearly as much as they might be and the reason undoubtedly lies in the lack of care taken in choosing a coupling resistor.

Integrating circuits

Circuits for integrating currents of the order of 10^{-12} amp. have a number of practical applications in the fields of radiation therapy and photometry. The standard electrometer circuits may be immediately employed for this type of work by the insertion across the grid circuit of a mica condenser of suitable capacity. Such circuits may be made to record automatically by employing a two tube circuit in which the output tube has a sharp plate current cut-off. A circuit of this type giving an accuracy of 0.1 percent has been described by Müller and Shriner.[26] For extremely small currents a circuit similar to that of Fig. 26 with the omission of the input resistor may be employed utilizing an FP-54 tube operated with electrometer characteristics.

Ballistic measurements

It is sometimes convenient to employ the plate current meter as a ballistic instrument and the circuit of Fig. 26 has been described for this purpose by Varian and Clark.[27] The ionization current is forced on to the collecting plate and after a suitable interval the key K_1 is closed, the resulting surge in plate current being recorded. If i_1 denote the current required to produce a unit deflection by the direct deflection method and i_2 the current required to produce the same deflection by ballistic measurement then if $rC \ll t$

$$i_1 = i_2 \cdot 2\pi T/t(1+\lambda/2),$$

where T is the time interval during which the current i_2 collects on the plate, t is the galvanometer period and λ the logarithmic decrement of the galvanometer. The ratio of ballistic to steady deflection sensitivity as a function of the grid impedance is a maximum when the time constant of the grid circuit is equal to the quarter period of the galvanometer.

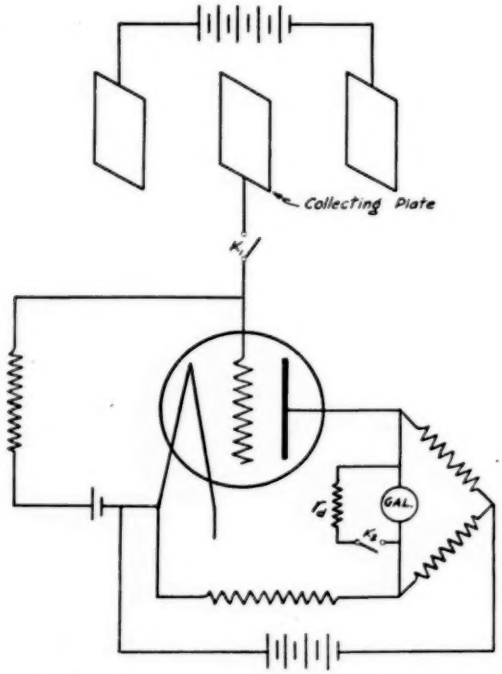

FIG. 26. Ions are driven on to the collecting plate and after a suitable time interval the accumulated charge is determined by connecting this plate to the control grid and reading the ballistic throw of the galvanometer. If the input resistor be omitted and electrometer characteristics employed for the valve, ballistic measurements are not necessary.

[26] R. H. Müller and G. E. Shriner, Rev. Sci. Inst. **6**, 16 (1935).

Another application of the ballistic principle has been made by Macdonald and Campbell[28] in a circuit especially developed for measuring radium emanation implants. With reference to Fig. 27 the radioactive unit is dropped into the ionization chamber I and caught at the center. The resulting surge of ionization current is amplified by the two tube resistance coupled circuit and recorded ballistically by the galvanometer G which is connected through a transformer to the plate circuit of the output tube. This circuit is of extreme value in this type of routine measurement as the galvanometer is unaffected by changes in the circuit potentials. The instrument has a limited linear response due to the properties of the iron core transformer.

Practical considerations

A practical limit to the voltage sensitivity of an amplifier is set by irregular movements of the plate current indicator and unless special precautions be taken these will probably be found to exceed the value computed from the shot and Johnston effects. The most common form of trouble is poor shielding and this is best overcome by placing the grid system in an aluminum

[27] R. H. Varian and J. C. Clark, Rev. Sci. Inst. **6**, 284 (1935).

[28] P. A. Macdonald and E. M. Campbell, Rev. Sci. Inst. **6**, 212 (1935).

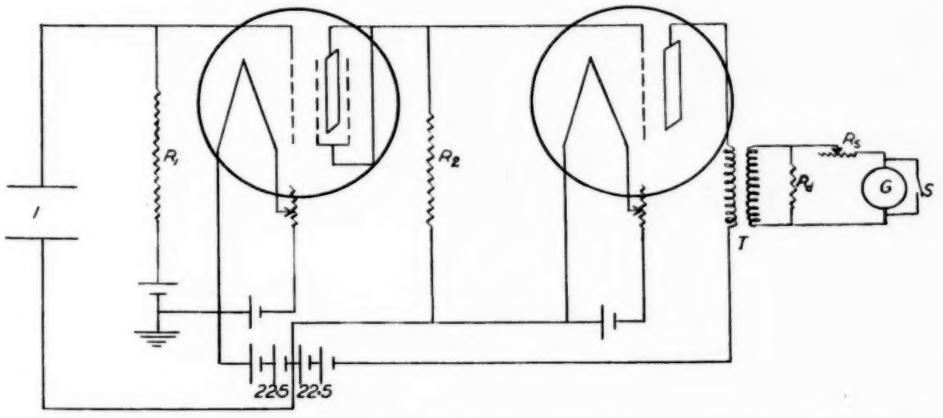

FIG. 27. Ballistic circuit in which the galvanometer is transformer coupled to the plate circuit of the output tube. This circuit is particularly insensitive to battery fluctuations and no thought need be given to its maintenance.

or copper box and then placing the whole amplifying unit, including batteries, in a second box. Both boxes should be connected to the negative side of the filament and then to ground by a soldered connection to a cold water pipe. When it is desired to place the batteries and galvanometer at some distance from the amplifier a double layer of braided copper wire shield may be pulled over the wire leads. Shields of this type may be secured from practically any dealer in radio parts. Poor electrical contacts, especially to storage battery terminals, often give rise to bad fluctuations and all contacts should be carefully soldered and in the case of storage batteries kept covered with clean vaseline. Poor soldering is worse than none and contacts to be soldered should be carefully cleaned, separately tinned, using a non-corrosive soldering paste, and then all traces of paste removed chemically. The final joining of the two tinned surfaces should be done with a clean iron, fresh solder and no paste. The General Electric Company manufactures a noncorrosive paste. Contacts made by the average potentiometer employed for radio construction are not satisfactory for voltage control and considerable care should be given in the selection of such equipment. The author has found that the 400 ohm potentiometer constructed by the General Radio Company is satisfactory for most work.

The selection of suitable tubes should be done with care. When currents less than 10^{-12} amp. are to be measured the highest degree of stability will be attained by employing the special direct current tubes. If expense is a major consideration radio tubes may be employed with satisfaction but they will seldom be devoid of drift, the best tubes showing a fairly constant drift of 10^{-6} volt/sec. The special tubes are thoroughly tested by the manufacturers and may all be relied upon but the author has found that on the average only 25 percent of the radio tubes of any one type are satisfactory. The stability of radio tubes may be tested by placing a sensitive galvanometer in the grid circuit and plotting the grid current change for a period of one hour after turning on the power, the filament and plate potentials being set at their rated values. After the first fifteen minutes good tubes will show a steady or slowly changing grid current while poor tubes will yield an irregular curve. No matter what type of tube is employed it will almost invariably be found that at least one-half hour will be required for the tube to reach equilibrium after the potentials have been applied.

In concluding these remarks it is to be emphasized that cleanliness is absolutely essential to stability. The tube and complete grid system must be cleaned to remove any trace of dirt especially grease and if handling is absolutely necessary it should only be done with rubber gloves.

Securing high resistance units at one time constituted a serious problem and a large number of different types of resistors have been sug-

TABLE IV. *Operating constants for direct current amplifiers.*

Current Range	Tube	Grid Circuit	Plate Circuit	Galvanometer	Operating Constants
10^{-15} amp. or less	FP 54	Electrometer Fig. 12	Fig. 7 or Fig. 20A	10^{-10} amp./mm	Fig. 21 and Table 2, col. 2
$10^{-12}\sim10^{-15}$	FP 54 or UX 222	Fig. 2B for FP 54 Fig. 16A for UX 222 or Fig. 12 at floating grid potential	Fig. 20A with FP 54 Fig. 7 with UX 222	10^{-10} amp./mm	Table 2, col. 1 for FP 54 Table I for UX 222
$10^{-9}\sim10^{-12}$	UX 222 or GE 232	Fig. 2B	Two tube circuit similar to Fig. 24 GE 230 as output tube	Microammeter	E_p, 18$\sim$22.5 $E_f=1.5$ GE 230 output, use E_p, 22.5 or 45 $E_f=1.5$ or 2.0
Greater than 10^{-9}	GE 230	Fig. 16	Fig. 16	Microammeter	$E_p=45$ $E_f=2.0$

gested from time to time. At present, however, satisfactory units are available in the open market. Resistances of 10^8 ohms or less may be secured from the International Resistance Company in units as large as 10^7 ohms. Values in excess of 10^8 ohms may be secured from the S. S. White Dental Company, New York or from the Westinghouse Electric Company whose resistors have been discussed by Rentschler and Henry.[29]

In addition to fluctuations arising from poor mechanical features, an irregular grid potential fluctuation of the order of 10^{-4} volt will be found due to the presence of air surrounding the grid system. The trouble will disappear if the tube and grid system be placed in a gas tight container and evacuated to a pressure somewhat less than 1 mm of mercury. The gas tight container is most conveniently constructed of brass. Leads through this container may be made by sealing brass rods into small bore Pyrex tubes with De Khotinsky cement, sealing the Pyrex to the brass with the same cement and then building up a layer of picein wax or Johnson canoe glue over the outside of the leads. The lid of the container need not be screwed in place but may fit loosely and then be sealed in place with picein or canoe glue provided both metal surfaces are heated above the melting point of the binder before its application.

In conclusion it is to be pointed out that the material in this paper has been selected mainly with a view to emphasizing the principles involved in direct current amplifiers rather than in giving detailed information about a number of specific circuits. Those who are interested in the immediate construction of an amplifier should find the information in Table IV of some assistance. In this table data are given for the construction of one particular circuit for each of the ranges of current that commonly arise in the laboratory.

The author wishes to express his thanks to the International Cancer Research Foundation, who generously supplied funds for a study of this subject with a view to developing some special types of photoelectric spectrophotometers.

[29] H. C. Rentschler and D. E. Henry, Rev. Sci. Inst. **3**, 91 (1932).

A Method for Studying Phenomena at the Cathode in Arc Discharges

L. R. KOLLER, *Research Laboratory, General Electric Company*

(Received May 25, 1936)

IN order to study cathode temperatures in hot cathode arc discharges, the following photo-tube method was developed. Essentially it consisted in operating the arc on a.c. and observing the photo-current due to the radiation from the cathode during the half-cycle that the arc was out. Because of the heat capacity of the cathode it did not cool appreciably during this time.

The type of arc studied is shown in Fig. 1. The cathode was an oxide-coated filament consisting of 6 turns of 20-mil tungsten overwound with $4\frac{1}{2}$-mil tungsten and coated with barium

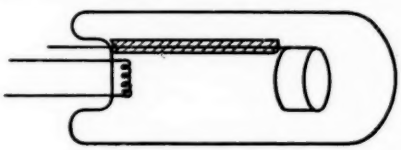

FIG. 1. Hot cathode arc.

hydroxide. The normal filament current was 10 amp. The anode was a cylinder of crimped molybdenum 1 inch in diameter and $\frac{1}{2}$ inch long. The distance between the cathode and the near edge of the anode was $1\frac{1}{4}$ inches. The gas pressure used was 1.5 mm neon. The photo-tube, the amplifier, and the arc were placed in a shielded box. The photo-current was amplified by a single stage of amplification using a UY 224 screen grid tube. Both photo-tube and amplifier were operated on a.c. 180° out of phase with the arc voltage. In order to avoid the possibility of the photo-tube recording any light from the arc which might persist into the "off" half-cycle, a d.c. negative bias was applied to the amplifier grid so that the plate current was suppressed except during the middle portion of the cycle.

The photo-current was measured at arc currents of 1, 2, 3, 4 and 5 amp. with the filament at 9, 10, 11 and 12 amp., respectively. A typical series of readings is shown in Fig. 2. The filament volt ampere characteristics are given below.

Filament amperes	Filament volts
9	1.0
10	1.3
11	1.55
12	1.94

At filament currents of 9 and 10 amp. the only effect of increasing arc current is an increase in cathode temperature. At 9 amp. filament current a 2.3 amp. arc will bring the filament to the same temperature as if it were operated at 10 amp. with no arc current. As the filament current is increased to 11 amp., a minimum in temperature is observed with increasing arc current. This is still more pronounced at 12 amp. This minimum is due to the cooling effect of the electron evaporation. The minimum is not observed at the lower filament temperatures because the filament is too cool to give the required arc current, and its temperature must be raised by ion bombardment to yield the electron current. At the higher filament currents, however, the filament is hot enough to supply the required current.

From these curves it is possible to make some estimates of phenomena in the arc. For example, the cooling effect of the electron evaporation can be obtained from the slope of the initial portion of the 12 amp. curve. The slope of the initial portion of this curve is 14 microamperes per ampere of arc current. Accordingly, the cooling effect of 1.75 amp. of arc current is 14×1.75 which is the equivalent of lowering the filament current from 12 amp. to 11 amp. with no arc current flowing. From the filament volt-ampere characteristic this is a decrease in wattage input of 6.2 watts. We may therefore put the heat lost by electron evaporation

$$n\phi e = 6.2\times10^{7} \text{ ergs}, \quad (1)$$

where n = number of electrons; e = electronic charge; ϕ = work function. Substituting numerical values

$$[1.75\times3\times10^{9}/e]\times(\phi/300)\times e = 6.2\times10^{7}, \quad (2)$$

$$\phi = (6.2\times10^{9})/(1.75\times10^{9}) = 3.54 \text{ volts}. \quad (3)$$

From the figure it can be seen that for a 12 amp. filament current the cathode is at the same temperature with a 4.7 amp. arc as with no arc current at all. Thus at this arc current the cooling by electron evaporation must just balance the

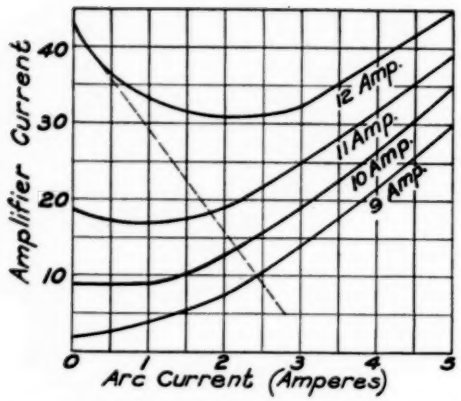

FIG. 2. Variation of photo-current with arc current at different values of filament current.

heating effect of the arc current. The latter is made up of the heating effect of the arc current returning through the filament and the heat due to the bombardment of the filament by the ion current.

The effective current is $4.7\times1.57=7.36$ amp. Accordingly, the i^2R loss in the filament due to the arc current is approximately $\frac{1}{3}(7.38^2)\times0.162=2.92$ watts. The heating due to the ion bombardment is

$$i(V+V_i)=i\times40 \tag{4}$$

where $V=$ tube drop and $V_i=$ ionization potential of neon. Therefore,

$$i\times40+2.92=4.7\times3.54, \tag{5}$$

and the positive ion current to the filament

$$i=(16.65-2.92)/40$$
$$=(13.73)/40=0.343 \text{ amp.} \tag{6}$$

Using the value of work function determined from Eq. (3) we can determine the positive ion current to the cathode under any conditions. For example, consider the point corresponding to 10 amp. filament current and 3 amp. arc current. The cooling due to electron evaporation is $3\times3.54=10.62$ watts. The heating effect of the arc current is $\frac{1}{3}(3\times1.57)^2\times0.13=0.95$ watts. From Fig. 2 it can be seen that under these conditions the filament is at the same temperature as if the current had been increased from 10 to 11 amp. with no arc current. From the filament volt-ampere characteristics this corresponds to an increase in energy input of 4 watts. Thus the combined effect of all the processes is to increase the filament temperature by the equivalent of 4 watts, or

$$i\times40+0.95-10.62=4$$

and

$$i=0.34 \text{ amp.}$$

These positive ion currents are of a different order of magnitude from the values which would be required if the arc current were supplied by field free emission from the cathode. For the case of a 3 ampere arc, if only a sufficient number of positive ions was required to neutralize space charge, the electron current would be 2.985 amp. and the ion current would be 0.015 amp. instead of the experimental value of 0.34. As long as the field free emission from the cathode is able to supply the required arc current, the arc current will always tend to cool the cathode. The heating effect as indicated by Fig. 2 is evidence that the cathode is only capable of supplying a small part of the arc current by field free emission.

This method is not capable of a high degree of accuracy since it can only be used with large values of arc current and under these conditions the heating effect of the arc current is considerable and difficult to correct for. The assumption is also made that the filament does not cool appreciably during the time of half a cycle.

The Coefficient of Thermal Expansion of Magnesium Oxide

MILO A. DURAND,[1] *Columbia University*

(Received June 12, 1936)

The thermal expansion of single crystal magnesium oxide has been measured from 85°K to 480°K. The results are combined with those of Austin, so that data are available from 85°K to 1273°K. The observed thermal expansion of this substance is in excellent accord with the relation, due to Grüneisen, between thermal expansion and heat content.

THE measurements reported in this paper were made with a vacuum interferometer precisely of the sort described in this journal by Austin.[2] Austin's experimental method was modified in two particulars: first, the temperature control was secured with a cryostatic bath instead of an electric furnace, and second, the temperatures were measured with a platinum resistance thermometer instead of a quartz refraction thermometer.

At temperatures below 273°K the cryostatic bath was that used by Balamuth[3] and Rose[4] in their apparatus for measuring the temperature variation of the elastic constants of solids. Above 273°K the bath liquid was pure glycerine, which was held in a large Dewar vessel, well stirred, and electrically heated. The temperature was stabilized to 0.1°K by manually controlling the current in an auxiliary heating coil of low heat capacity. The two baths were placed in succession on an elevator, and raised until the interferometer was properly immersed. Thus observations were made continuously over the entire temperature range.

The platinum resistance thermometer was calibrated by the U. S. Bureau of Standards, and was used in connection with a Leeds and Northrup Müller bridge.

The specimen material was obtained from large, clear, well formed crystals furnished by the Norton Company of Worcester, Mass.

[1] Now with Consolidated Film Industries, New Jersey.
[2] Austin, Physics **3**, 240 (1932).
[3] Balamuth, Phys. Rev. **45**, 715 (1934).
[4] Rose, Phys. Rev. **49**, 50 (1936).

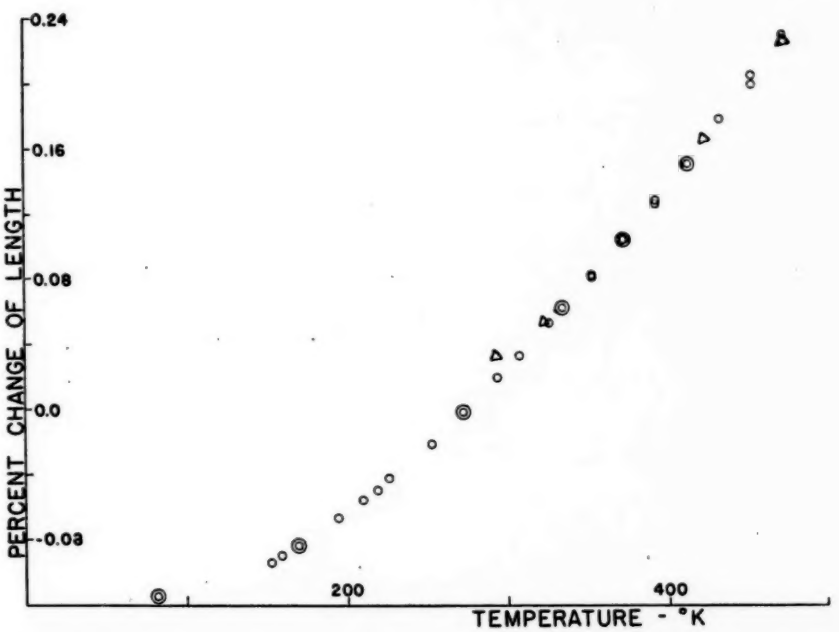

FIG. 1. The thermal expansion of magnesium oxide. The circles represent the present data and the triangles those of Austin fitted at 373°K.

TABLE I. *The thermal expansion of magnesium oxide.* (The data above 480°K are Austin's)

Temp. (°K)	Total expansion in percent of the length at 273°K	Temp. (°K)	Total expansion in percent of the length at 273°K
0	−0.116*	390	+0.129
85	.114	400	.141
150	.092	410	.153
160	.087	420	.166
170	.081	430	.178
180	.075	440	.191
190	.069	450	.204
200	.061	460	.216
210	.053	470	.229
220	.045	480	.242
230	.037	523	.297
240	.028	573	.362
250	.020	623	.432
260	.012	673	.499
270	.002	723	.572
280	+0.008	773	.644
290	.019	823	.717
300	.029	873	.789
310	.039	923	.862
320	.049	973	.937
330	.060	1023	1.012
340	.071	1073	1.087
350	.082	1123	1.162
360	.093	1173	1.237
370	.105	1223	1.312
380	.117	1273	1.387

* This value was obtained by extrapolation with Grüneisen's formula, q.v. infra.

Observations were made on two sets of specimens, one about 3 mm in length and the other about 5 mm. The results obtained with the two were indistinguishable from each other.

The total expansion of magnesium oxide, expressed in percent of the length at 0°C, is shown graphically by the circles in Fig. 1. The triangles on the same figure indicate the data on this substance reported by Austin.[5] These are plotted in such manner as to force agreement with the present observations at 373°K. There seems to be a slight inaccuracy in the two observations of Austin below this temperature, while above it the two sets of data are in excellent accord.

The present data between 85°K and 480°K are given in Table I, together with Austin's results between 473°K and 1273°K. As in the figure, the latter have been forced into agreement with the former at 373°K.

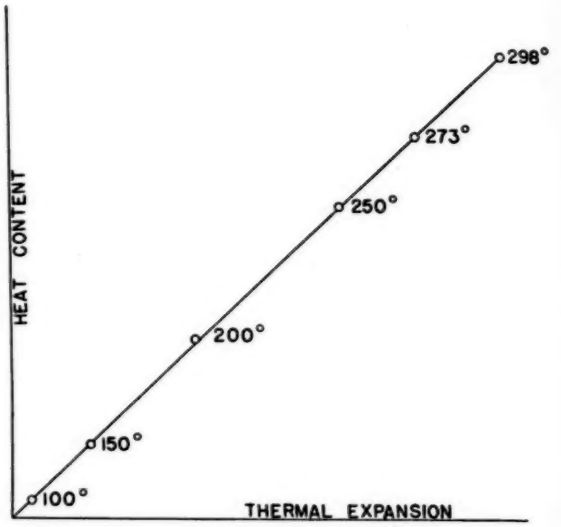

FIG. 2. In accordance with Grüneisen's relation, the thermal expansion of magnesium oxide varies linearly with the heat content.

It is of interest to note that the observed behavior of magnesium oxide is described very exactly by a simplified relation between thermal expansion and heat content obtained by Grüneisen[6] from the lattice theory of crystals. This may be written in the form $\Delta l/l_0 = c\Delta H$, where l_0 is the length at absolute zero, Δl and ΔH are the changes in length and heat content, respectively, and c does not depend on the temperature. The actual linear relationship between Δl and ΔH is shown in Fig. 2. The indicated values of ΔH are given in the *International Critical Tables*.[7] Unfortunately this quantity has not been measured above 298°K, which is the temperature corresponding to the highest point shown on the graph.

The author gratefully acknowledges his indebtedness to Mr. Raymond R. Ridgway, of the Norton Company, who was instrumental in securing the specimen material, and to Dr. S. L. Quimby, under whose direction this work was done as a part of a larger program of research.

[5] Austin, J. Am. Ceramic Soc. **14**, 795 (1931).

[6] Grüneisen, *Handbuch der Physik*, Volume X, p. 42.

[7] *International Critical Tables*, Volume V, p. 91.

A High Frequency Apparatus for Producing a Measured Heat Stimulus in Studies of Physiological Temperature Regulation[1]

ALLAN HEMINGWAY AND GLENN H. WITTS, *The Laboratory of Physiological Chemistry, University of Minnesota*

(Received June 13, 1936)

The requirements of a high frequency diathermy machine for clinical therapy and for studies of physiological temperature regulation in regard to (1) frequency, (2) measurement of power output, (3) range of power output, (4) efficiency and construction are discussed. To conform with the above requirements a power oscillator of 1000 kilocycles has been built capable of delivering up to 200 watts of high frequency power. A description of the apparatus is given with circuit constants and meter values which have been found useful in physiological temperature studies on animals and human subjects. With this apparatus it is possible to measure the power consumption of a subject treated with diathermy from the voltmeter and ammeter readings.

IN studies of temperature regulation it is often desired to administer to an animal or a human subject a known quantity of heat. With the usual methods of producing fever an animal is placed in a warm bath or a warm air environment or exposed to the rays from an infrared lamp. With these methods the quantity of heat received by the animal cannot be measured and they cause a certain amount of discomfort due to causes other than heat. The heating of an animal by high frequency currents (diathermy) offers a method which is free from discomfort to such an extent that the animal is unaware of the passage of the current and feels only the generated heat. The measurement of the heat production can be made in a very simple manner by measuring the voltage and current with high frequency meters since it has been shown by Hemingway and McClendon[2] that with a proper choice of frequency the phase angle of a tissue is zero.

REQUIREMENTS OF DIATHERMY MACHINE[3]

Frequency. In order to be used for heating tissues which can be stimulated electrically it is necessary to use alternating currents of a frequency greater than 500 kilocycles per second to prevent neuromuscular stimulation. For measurements of the heat production in tissues by the ammeter-voltmeter method, in which the phase angle is zero, the frequency must remain within a range of 500 to 2000 kilocycles.[4] As the most suitable frequency which would satisfy both of these requirements a frequency of 1000 kilocycles per second is the most desirable.

Power output. It has been shown that the rate of heat production in patients for therapeutic purposes varies from 20 to 200 watts[4] with most of the routine treatments being of the order of 50 to 100 watts. The machine described in the paper has been designed to have a maximum power output of 200 watts.

Construction and insulation. Where high 60-cycle or d.c. voltages are used in the high frequency generator, especial precautions must be taken to prevent the subject heated from coming into contact with a high voltage which might cause an electrical shock. The safest method of fulfilling this requirement is to place the subject in a tuned circuit inductively coupled to the main oscillator. In most of the usual spark gap diathermy machines now used clinically the electrodes to the patient are attached to the inductance coil of the main oscillator, the patient being protected from the stimulating voltages by a blocking condenser.

Regulation of output current. Due to the wide variation of output power it is necessary to have some device by which the current through the patient can be carefully, quickly and simply regulated. This factor is especially important when one considers that the upper limit of the diathermy current which can be tolerated is determined by the current density of the smallest

[1] Aided by a grant from the American Medical Association.

[2] Hemingway and McClendon, Physics **4**, 351 (1933).

[3] Hemingway, J. Am. Med. Assoc. **101**, 776 (1933).

[4] Hemingway, Radiology **14**, 99 (1930).

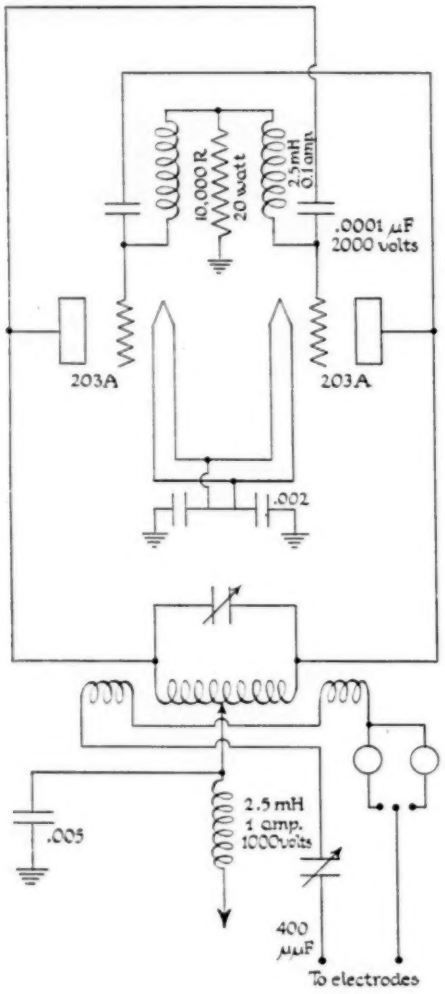

FIG. 1. Circuit diagram of diathermy machine.

diathermy electrode;[5] hence the current must be adjusted for the area of the smallest electrode and since this area is quite variable a sensitive regulation is necessary.

Efficiency. It is desirable for purposes of economy of operation and for prevention of over heating of the apparatus to have a fairly high "over all" efficiency, defined by the ratio of power output to power input. The efficiency for most diathermy machines varies from 20 to 60 percent.

CHOICE OF CIRCUIT

The circuit which was finally used after some preliminary experimentation is shown in Fig. 1. This is of the standard two tube power oscillator type with two 203A tubes. The resistor, capacity and choke coil values are given on the diagram and were those which were found to give the most efficient operation. The subject is heated by the high frequency current passing between two electrodes applied to the skin. As shown in Fig. 1, the subject heated is in a tuned circuit inductively coupled to the inductance of the main oscillator. Two coupling coils are situated at the ends of the inductance coil of the main oscillator. In series with these two coils is a variable air condenser and a high frequency ammeter. The high frequency voltmeters are connected by plug connections to the leads which go to the electrodes. The power supply unit is not shown.

This machine satisfied the above mentioned requirements. Most diathermy machines used in clinical heat treatments are of the spark gap type. An advantage of the vacuum tube oscillator type of machine, shown in Fig. 1, is that the peak voltage is much lower than that of the spark gap machine. The high spark gap peak voltage causes arcing across the symmetrical shields[6, 7] of the high frequency voltmeter which does not occur with the continuous sine wave voltage from the vacuum tube oscillator.

CONSTRUCTION OF DIATHERMY MACHINE

The construction of the machine is shown in Fig. 2. An iron framework of light angle iron supports two shelves. On the upper shelf is placed the oscillator tubes, a 240 $\mu\mu$F variable air condenser with capacity fixed to give 1000 kilocycles and the inductance coil of the main oscillating circuit. The high frequency choke coils and resistors are attached to the lower surface of the upper shelf. The plates of the Cardwell condenser must be spaced 0.1 inch apart and heavy copper wire used in the inductance coil. The coupling coils are attached to Bakelite rods which can be clamped at variable distances from the inductance of the main oscillating circuit. A safe distance is one inch which ensures all the insulation required between the high plate voltage (1000 volts d.c.) and the patient circuit. The variable air condenser (440 $\mu\mu$F, space between

[5] Hemingway and Stenstrom, J. Am. Med. Assoc. **98**, 1446 (1932).

[6] Wilson, Trans. A.I.E.E. **43**, 220 (1924).

[7] Goodwin, Trans. A.I.E.E. **46**, 479 (1927).

plates 0.07 inch) is attached to the roof of the masonite cabinet which encloses the iron frame. The ammeters are Weston radiofrequency ammeters one having a full scale deflection of 1.0 ampere and the other 2.5 amperes. The voltmeters are Weston high frequency thermocouple voltmeters, one having a full scale deflection of 125 volts and the other 50 volts. The voltmeters can be attached by plugs to sockets on the panel. Two power supply switches with corresponding pilot lights are used. One switch is in the main 60-cycle line and filament circuit while the other switch is in the plate circuit. By closing the switches in the proper order the application of a high voltage between the cold cathode and plate is prevented.

Operation

In using the machine electrodes are placed on the skin at the place where the heat is to be applied. The filament switch is first turned on followed about two minutes later by the plate switch. The current through the subject is raised to the desired value by increasing the capacity of the variable condenser in the patient circuit. An accurate adjustment is possible with this type of regulation. The heat production in watts is simply the high frequency current multiplied by the high frequency voltage since under these conditions the phase angle is zero and the thermocouple meters indicate r.m.s. values.

Results of Tests

This apparatus has been tested on humans and dogs. There is no evidence of electrical stimulation. The power output can be varied from 0 to 200 watts by the variable tuning condenser. It is advisable to delay connecting the high frequency voltmeters until the load is in the circuit, otherwise the high voltage throws the pointer off scale.

Copper-advance thermocouples have been used to measure skin and rectal temperatures

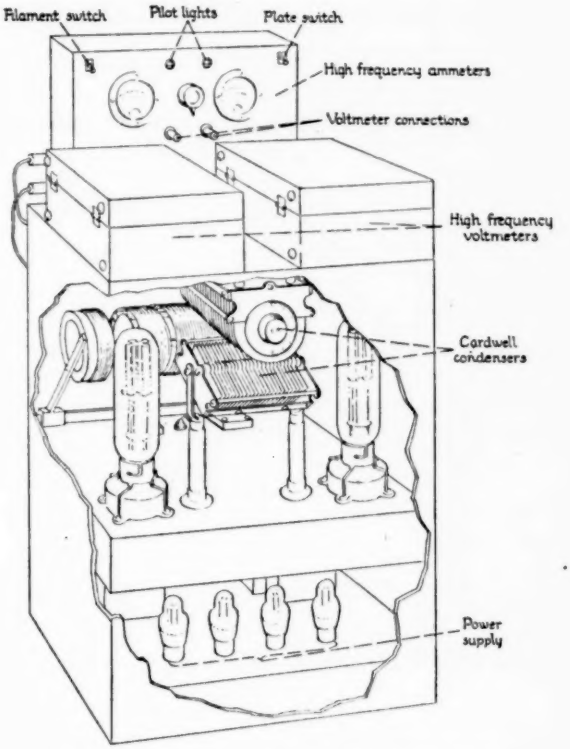

Fig. 2. Diathermy machine showing construction.

while the machine is in operation. The high frequency current is without effect on the thermocouple voltage measurement if high frequency choke coils are placed in the thermocouple galvanometer circuit and the bimetallic junctions are covered with a thin layer of vaseline. Without the vaseline covering the stray field of the oscillator induces a small high frequency current in the thin thermocouple wires which causes a local heating where the bimetallic junction makes contact with the skin. The error involved is about 0.5°C.

The authors wish to express their appreciation to Dr. H. E. Hartig for his assistance in some of the technical difficulties involved in the operation of the machine and Mr. H. A. Carter and the Council of Physical Therapy of the American Medical Association for financial aid.

RHEOLOGY SECTION

Further Tests of the Elasticity of Protoplasm[1]

HANS H. PFEIFFER,[2] *Bremen, Germany*

(Received January 24, 1936)

It is possible to produce changes in fluidity of protoplasts by means of the varying shearing stresses to which they are subjected. In consequence of the different fluidity of the objects, the speed of their deformation is dependent upon the pressure or shearing stress used. In further experiments the relative effects of changes of temperature and of different concentrations of anisotonic media were studied, and in supplementary experiments with inanimate model drops, it was verified that the phenomenon of increasing "apparent fluidity" depending on the shearing stress is probably due to the structure causing elasticity. Not only the plasmalemma, but also the interior of the protoplasts is more fluid as the shearing stress increases, for the Brownian movement of particles of the inner layer is more rapid. From these experimentally determined facts the conclusion is drawn that the tangential expansion for protoplasm, as of any non-Newtonian fluid, is a function of the tangential shear γ and its change in time:

$$\tau = f(\gamma \cdot d\gamma/dt).$$

INTRODUCTION

1.

As a rule, protoplasm appears to be a fluid, but the microdissection experiments of Seifriz[3] and Scarth[4] seem to contradict this. We must, however, distinguish between true elasticity and a mere dilatability or deformability. We should not deduce elasticity from measurements of the limit of extensibility or of the ability to draw threads.[5] Also, osmotic increase of protoplasmic volume is not of use in the measurement of elasticity. As is well known, a surface film can behave like an elastic membrane. Moreover, even in mass, solutions which we ordinarily regard as fluid may occasionally show form elasticity. For instance, a gelatin sol may have elastic properties in spite of the fact that it flows fairly easily, for the fluidity (or viscosity) of the gelatin sol is a function of the pressure or shearing force to which it is subjected. We measure elasticity by means of the electromagnetic attraction of a minute nickel particle embedded in the protoplasm.[6]

If protoplasm is elastic, it cannot be a true liquid. Consequently Stokes' formula is not applicable, and, therefore, neither the gravity nor the centrifuge methods are suitable. Hence, we do not regard as conclusive L. V. Heilbrunn's centrifuge experiments with varied rotatory speeds.[7] A much better argument would be a demonstration of anomalous flow of protoplasm; i.e., a higher fluidity at increasing mechanical shear. But unfortunately we cannot cause the protoplasm to flow through a capillary. So I use another method which has been employed by Fr. Vlès[8] with eggs of tardigrades. Also I have combined the method with J. Pekarek's[9] measurement of the absolute viscosity by determining the number of the "mean two-way first passages"* of particles in Brownian movement.

[1] This paper has been reported in part at the 6th Internat. Botanic Congress in Amsterdam, but the subject matter partly exceeds the statements given there.

[2] Department of Botany, Kolonial-und Ueberssee-Museum, Bremen.

[3] Wm. Seifriz, Brit. J. Exp. Biol. **1**, 431 (1924); Protopl. **1**, 345 (1926) and **9**, 177 (1930); J. Rheology **1**, 261 (1930); in collaboration with Janet Plowe, J. Rheology **2**, 263 (1931).

[4] G. W. Scarth, Protopl. **2**, 189 (1927).

[5] H. Pfeiffer, Kolloid-Zeits. **70**, 26 (1935).

[6] A. Heilbronn, Jahrbuch wiss. Bot. **61**, 284 (1922); H. Freundlich and W. Seifriz, Zeits. f. physik. Chemie **104**, 233 (1923).

[7] L. V. Heilbrunn, J. Exp. Zool. **43**, 313 (1926).

[8] Fr. Vlès, Bull. Inst. Océanogr. Monaco, No. 589 (1931); Arch. zool. exp. et gén. **75**, 421 (1933).

[9] J. Pekarek, Protopl. **10**, 510 (1930); **13**, 637 (1931); **17**, 1 (1932), and other volumes.

* (If *a*, *b*, *c*—denote successive parallel scratches on a microscope scale, Brownian particles in their zigzag movement will make *a passage* across some line, say *b*. If the particle crosses the line *b* again immediately, this successive passing does not count. In one-way passings, the next passage must be across the line *c* and the third across

Experimental Procedures and Results

2.

The naked cells used were obtained partly from liquefying bacca pericarps (Physalis Alkekengi, some Solanum species, Juniperus communis, Vitis vinifera, etc.), partly by means of experimental denudation of the protoplasts[10] within epidermal cells (onion-skin of Allium Cepa) and parenchyma cells (from leaves of Hyacinthus orientalis). With the behavior of these objects I have compared cells present in saliva (resembling leucocytes) obtained from the human mouth.

The apparatus of deformation (Fig. 1) consists of a capillary c, a water manometer m, and a tube v with india-rubber aspirator bulb and two-passage tap, all three united by a Y tube z between.[11] After sucking up the protoplasm by means of a vacuum, measured on the attached manometer, we observe and measure the deformation of the objects.

3.

The deformation of the protoplasts requires time. We measure the time necessary for filling a precise length of the capillary (in cm 10^{-4}) at any negative pressure (in cm H_2O). The curve of deformation speeds v (in cm 10^{-4}/sec.) plotted against the vacuum head p exerted on the objects shows a remarkable peculiarity

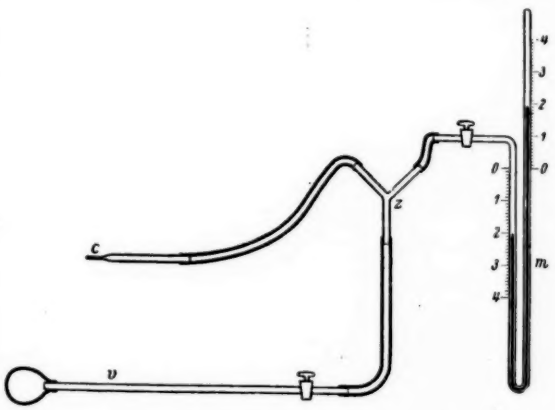

Fig. 1. Apparatus for inducing graded shearing forces.

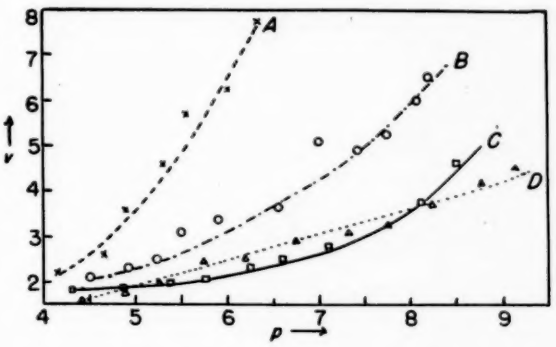

Fig. 2. Curves of deformation speed v plotted against the shearing force p demonstrated on *protoplasmatic drops* of Juniperus pericarps (A–C) after plasmorrhysis in 0.5 (x - - - x), 0.6 (o - - - o), and 0.65 g (□——□), moreover (D) on *inanimate droplets* of glycerin (△. . . . . △).

(Fig. 2). It is not a straight line, as it would be expected if protoplasm were an ideal liquid. The quotient of p to speed v is not constant, but the velocity of the shearing flow increases with increasing force, at first slowly, later more and more quickly.

The surprising analogy of these curves with those of the anomalous flow of colloidal substances is shown far better if we regard the influence of degree of plasmorrhysis or of the experimental temperature.

If hypertonic media of different concentrations (salts or nonelectrolytes) act upon the naked protoplasts, we observe their plasmorrhysis; i.e., wrinkling and shrinking in volume, according to the concentration of the medium. These plasmorrhyzed cells also differ from normal ones in respect to their behavior toward the shearing stresses in our apparatus, for the slope of our curve of deformation-speeds decreases, the more the plasmorrhysis rate increases (Fig. 2). Chiefly to this circumstance I attribute an often considerable scattering of the values obtained from liquefying pericarps.

On the other hand, the slope of the curves decreases as the working temperature either increases to 25 to 30°C or decreases to 12 to 15°C. Here we find a relation similar to that of temperature on the viscosity of protoplasm.

I have compared our curves with those of model liquids, such as water, glycerol, sodium stearate, etc. In these experiments, the values are nearly on a straight line if the fluid is inelastic, as in water or glycerol. But there is obtained

d, etc.; but in two-way counting of the first passings, the crossing of b may be followed by a crossing of either a or c, hence the measurement of the mean time required for "first passages" according to the "two-way method" of counting.—*Editor.*)

[10] H. Pfeiffer, Zeits. wiss. Mikrosk. **50**, 429 (1934).

[11] H. Pfeiffer, Protopl. **23**, 210 (1935).

quite an appreciable curve if the material used shows anomalous flow, as with sodium stearate.

4.

Whether our curves are comparable to such an anomalous flow of elastic fluids, however, cannot be decided with certainty without measuring the differences in viscosity of the material before and after shearing. For this purpose, I use J. Pekarek's method. There I find the absolute values of the viscosity by means of the determination of the number of the mean two-way first passages of particles in Brownian movement from

$$\bar{\eta}=\frac{R}{N}T\cdot\frac{t}{3\pi al^2n}=\frac{R}{3\pi l^2N}\cdot\frac{Tt}{an}$$

(where R=absolute gas constant, N=Loschmidt's number, T=absolute temperature, t=time during the counted passages, a=radius of the particle, l=distance between two scale lines of the ocular, n=number of the passings).

Using this method with the naked protoplasts we prefer objects with a relatively thick cytoplasmic surface layer. Then we do not need to introduce R. Ladenburg's correction, and, moreover, we become certain of observing particles wholly within the cytoplasm. If we do not use the most favorable cells, the applicable material, it is true, is limited. We obtain useful material by opening the cells of *Chara fragilis* and staining them *intra-vitam* with 0.05 percent chrysoidin. We do not need to inquire whether the colored particles are preexistent inclusions (microsomes, droplets of different substances) or new formations. It must satisfy us that they are in action Brownian movement.

In order to have a measure of comparison, we assume the value of the absolute viscosity (determined without a shearing force) ($\bar{\eta}\sim0.0498$) as 100 percent. Then we measure the viscosity values through a whole series of increasing heads and connect the values with the initial one. The values of the viscosity, as shown on the coordinate scale, are distinctly lower as the rate of shear increases (Fig. 3). The rate of flow is not proportional to the increasing shearing force as expected for a true fluid. Under suction the viscosity tends to be less, and as the shearing stress is increased, there is an increase in speed of the Brownian movement.

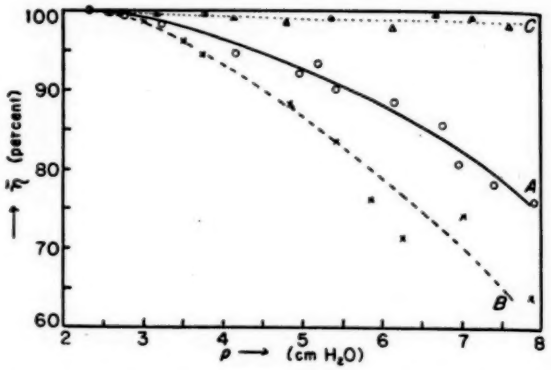

FIG. 3. The absolute viscosity after influence of shearing forces expressed as percentage of the initial viscosity plotted against the shearing force demonstrated on *protoplasmatic drops* of cutting cells of Chara fragilis (*A–B*) at a temperature of 21° (o——o) and 12°C (x - - - x), moreover (*C*) on *droplets of glycerin* at 21°C (△ . . . . . △).

Taking into consideration the temperature, we find increasing fluidities between 26° and 31° just as between 10° and 15°C. The influence of increasing plasmorrhytic rates has not been examined sufficiently as yet, so that the results are not conclusive. But model experiments with water, glycerol, or sodium stearate show after vital staining a distinct difference between true fluids, such as water or glycerol, the flowing of which is proportional to the shearing stress, and elastic fluids which have a lower viscosity as the rate of shear increases.

5.

The ability of drops to resist distortion must depend in part on the surface tension of the film surrounding the naked protoplasts, it must also depend on the elasticity of the plasmalemma (plasma membrane). Whether the interior of the protoplast is elastic or not is a problem which requires further study than the mere observation of the elastic properties of the object as a whole. It has been shown that the fluidity of the naked protoplasts is a function of the pressure or shearing stress to which it is subjected. Not only does the fluidity of the plasmalemma increase with the suction, but also the interior of the protoplasts is more fluid as the shearing stress increases so that it is thixotropic. This we conclude from the increasing speed of Brownian movement of particles of the inner protoplasm.

Usually a plastic solid or an elastic solid owes its elasticity to the presence of structural elements. Under pressure or distortion, the structure tends to break down, and this is reflected in a pronounced decrease in viscosity. When elastic materials are subjected to a shearing stress, they are usually double-refractive. Researches on cells under crossed Nicol prisms have shown, now and then, such double-refraction, but in other cases one does not observe marked anisotropy. Further investigations are, therefore, likely to yield results of considerable interest.

Plastic Properties of Shellac[1]

WM. HOWLETT GARDNER,[2] *Polytechnic Institute of Brooklyn, Brooklyn, N. Y.*

(Received November 26, 1935)

The plastic properties of shellac is a subject of international interest, but one which has received relatively little study. In this paper, those properties such as polymerization, aggregation, condensation in the presence of chemicals, and adsorption of water, which would affect rheological measurements are reviewed and discussed.

MANY interesting rheological problems for study are suggested by the physical and chemical characteristics of lac, since this resin is quite unusual in containing both permanently thermoplastic, and thermosetting constituents.[3] The mean size of the molecules is such as to impart both colloidal and crystalloidal properties,[4] so that the phenomena of colloidal aggregation, and of molecular condensation-polymerization are observed on thermal treatment.

Lack of Rheological Data

In spite of the fact that these general plastic properties have been known for a long time, there has been but one attempt to obtain actual rheological values.[5] This is indeed surprising when it is recalled that the practical applications of shellac depend upon either its plastic fluidity[6] or upon the viscosity of its varnishes. In certain industrial uses both properties are significant.

The reasons for this lacuna in our knowledge of its physical properties are to some extent understandable from a technical standpoint. Shellac has been used from the very beginning of the commercial development of modern plastics,[7, 8] and the study of its characteristics has to a large extent been carried out on the production floor rather than in the laboratory. Since shellac is too brittle to be used by itself except in thin layers, individual users have preferred methods of trial and error for the study of combinations, and have learned in this manner to obtain and control the desired working properties by either varying the proportions of the ingredients in their compositions, or by blending different grades of lac. The fact that shellac was used by many concerns when they first organized has unfortunately not added to the incentive to undertake these much needed scientific studies.

New Interest in Rheology

This picture, however, somewhat changed during the past year. Electrical manufacturers in both the United States and Great Britain are now taking a renewed interest in this resin with the object of learning more concerning it. This is especially true with respect to fluidity, where even a preliminary study of control methods has given important information as to how manufacturing processes may be improved.

This revival of interest has probably in part been stimulated by the recent international discussions on specifications and test methods of lac,[9, 10] and in part by the publicity which shellac research received when it was expanded and internationalized, through cooperative arrangements between the different research centers.[11]

[1] Presented at the Plastics Symposium of the Seventh Annual Meeting of the Society of Rheology at the Bell Telephone Laboratories, New York, N. Y., Oct. 12, 1935.

[2] Research Professor of Chemical Engineering and Supervisor of Shellac Investigations.

[3] W. H. Gardner and B. Gross, British Plastics **6**, 514 (1935).

[4] W. H. Gardner, B. P. Caldwell and H. Burnett, "Research Note on Molecular Weight Determinations on Shellac," in press.

[5] R. Houwink, *Physikalische Eigenschaften und Feinbau von Natur und Kunstharzen* (Akad. Verlag, Leipzig, 1934); *cf.* L. C. Verman, London Shellac Res. Bur., Tech. Paper No. 4 (1935).

[6] Possibly this property should more strictly be called mobility but our knowledge is so meager that it is impossible to tell whether we are dealing with a non-Newtonian liquid or with a plastic. The author has therefore chosen to use the term fluidity in a general sense throughout this paper.

[7] A. J. Weith and A. V. H. Mory, Am. Inst. Chem. Eng. Silver Ann. Vol., 213 (1933).

[8] W. H. Gardner, British Plastics **6**, 459 (1935).

[9] H. L. Curtis, Proc. Am. Soc. Testing Materials **34**, 502 (1934).

[10] Report of Committee D-9, Am. Soc. Testing Materials (1935).

[11] Chemical Industries **36**, 167 (1935).

The present is therefore a very auspicious time to review those properties of lac which affect and influence rheological measurements.

Polymerization

From analysis, pure shellac, aside from small amounts of wax and other natural impurities, is a solid solution of similar resins, which are intra-esters of poly-hydroxy acids.[12, 13] These resins have both uncombined carboxyl and hydroxyl groups free for further condensation under heat, and are probably composed of three or four different acids which are either aliphatic[14, 15] or benzoid[16] types. All the lac resins show marked adsorptive properties for both solids and liquids.

When shellac or its compositions are subjected to prolonged heating, definite changes take place. At first there is not a great deal of change in fluidity and melting point, but towards the end of the period of "life" or "time for polymerization" at each temperature there is a rapid increase in both these values.[17, 18] This change is immediately followed by a transformation to a more or less rubbery consistency which produces a tough, brittle, horny mass when cooled.[19]

During the polymerization there is a small but steady loss in water, which may vary up to 5 percent of the weight of the lac. No other volatile material is evolved.During this process increasing proportions of the resin are rendered insoluble in alcohol,[20, 21] but with temperatures up to the decomposition point of lac (about 300°C) about 5 to 10 percent will always remain in a thermoplastic and soluble state. These thermoplastic resins are diethyl ether soluble and act as a natural plasticizer for the rest of the material. When polymerized lac or "cured" compositions are molded under high heat and pressure this plastic resin flows to the surface and imparts a gloss or polish to the object when it is cooled.[17] The presence of such resin constituents would account for the fact that the usual shellac of commerce never completely exhibits true thermosetting properties,[3] but that lac purified by extraction with ether or other solvents shows evidence of resinoid characteristics.

In considering the polymerization of lac, it may prove convenient to borrow thermosetting terminology and to designate as *A* to *C* the three distinct types of resinous materials obtained.[3] Lac which has not been heated to lose its flow or softening point would be classified as "A" stage lac. This lac will produce lac "B" by condensation-polymerization through chemical loss of water between hydroxyl and acid groups of different molecules in forming esters, or between two acid groups of different molecules in an anhydride formation. It is unlikely, as the amount of water split off would indicate, that more than a two or three-fold increase in molecular weight takes place during this condensation. The sudden change in physical properties which occurs toward the end of the condensation or so-called "life period" probably represents a production of a gel structure in much the same manner as has been shown to take place in the drying of linseed oil,[22] or in the heating of alkyd resins.[23] Hardening of this gel will give lac "C."

"B" lac, while it does not show any appreciable fluidity or softening at atmospheric pressures, can still be molded at temperatures of 130–160°C (270–325°F), under pressures of 900–1500 g/cm² (1800–3000 lb. per sq.ft.). The three-dimensional space arrangement of the molecules of lac "B" is probably quite complex. Further heat treatment, while it may not increase the molecular weight through condensation, may cause the molecules to become more closely entangled and to assume a more compact structure, which would account for the properties of the final stage or lac "C." This resin is very brittle and tends to crumble under stress of shear with further heat and pressure.

---

[12] W. F. Whitmore, H. Weinberger and W. H. Gardner, Ind. Eng. Chem., Anal. Ed. **4**, 48 (1932).
[13] R. Bhattacharya, J. Soc. Chem. Ind. **54**, 82T (1935).
[14] C. Harries and W. Nagel, Ber. **B55**, 3833 (1922).
[15] W. Nagel and M. Körnchen, Wiss. Veroeffent. Siemens Konz. **6**, 235 (1927).
[16] W. H. Gardner, G. Pribyl and H. Weinberger, Ind. Eng. Chem., Anal. Ed. **6**, 259 (1934).
[17] W. H. Gardner, B. Gross, C. C. Whipple and M. Fasig, British Plastics **6**, 571 (1935).
[18] B. Gross and W. H. Gardner, unpublished research note.
[19] S. Ranganathan and R. W. Aldis, Indian Lac Research Inst. Bull. No. 14 (1933).
[20] W. Nagel, Metallborse **20**, 2133, 2190 (1930).
[21] H. Stager, *Kolloidchemische Technologie* (Th. Steinkopff, Dresden und Leipzig, 1927), p. 292.
[22] J. S. Long, A. E. Rheineck and G. L. Ball, Jr., Ind. Eng. Chem. **25**, 1086 (1933).
[23] R. H. Kienle, Ind. Eng. Chem. **23**, 1260 (1931).

Of course this picture of a gel structure is to some extent pure conjecture. It certainly would be of great interest to verify or disprove by detailed experimental measurements these theoretical concepts, since it would not only greatly add to our knowledge of physical properties of resins in general, but also a considerable amount to our knowledge of electrical insulation. What has been termed "cured lac" in this latter field is still a very vague term even from a manufacturing angle.

Aggregation

The polymerization of lac through the chemical splitting off of water which takes place when shellac is heated should be distinguished from the aggregation, described by Harries and Nagel[24], which results when shellac is treated, as for example, with ether containing a little hydrogen chloride or when it is stored for a long period of years.[25, 26]

Both processes yield products which are insoluble in alcohol and are more or less infusible,[20] but their chemical and physico-chemical properties are distinctly different.[15] The loss of water in the former case changes the chemical composition, while with the latter there is no change in empirical formula.[14] The polymerized material resulting from aggregation has a decreased saponification number and can not be completely broken down into its constituent acids by means of caustic. Heat polymerized lac has the same saponification value as the original resin. The aggregated material may be depolymerized by successive solution and precipitation from first acetic acid and then alcohol,[15] while the heat hardened material must be depolymerized by steam and pressure.[27] This latter reaction confirms the observation that the heat hardened material is a condensation product at least in its first stage. Depolymerization of this resin by acidulated alcohol[28] can be explained upon the same basis. In this case it is probably a combination of hydrolysis and alcoholysis. From an industrial standpoint depolymerization in an autoclave with steam is naturally to be preferred to depolymerization with alcohol. Neither process is as yet used commercially.

Condensation in the Presence of Chemicals

When shellac is heated with poly-reactive chemicals capable of reacting with one or more of the groups present in the lac, accelerated polymerization usually takes place. In this manner it is possible commercially to shorten to any desired extent the life of a given lac.[29] A wide variety of chemicals can be used, such as polycarboxylic acids or their anhydrides, polyamines, dicyandiamide, biguanides, or oxides of poly-valent metals, as lime. Strong acids, aldehydes and ketones also produce similar effects. However, other mono-reactive substances are definitely limited in their effect.

The chemical behavior of the poly-reactive types is readily explainable from our knowledge of synthetic resins. The reactions with shellac are however more complex than those in the usual condensation, since reactions can take place between individual shellac molecules simultaneously with reactions between the lac and the added chemical. Hence, the amount of reagent which is chemically incorporated in the resulting product is not a stoichiometrical quantity. To some extent it is dependent upon the original proportions used and upon the conditions during reaction. In certain instances characteristic properties are imparted to the product, which are not usually obtained by heat condensation. For example, while with most of the poly-reactive chemicals there is a very slow initial increase in fluidity accompanied by a gradual decrease in alcohol solubility, with phenylenediamine there is a very rapid increase in the viscosity of the melt and little or no decrease in solubility during the preliminary stages. It is possible that during condensation some chemicals can favor certain individual reactions at the expense of other types.

Poly-reactive acids probably play a dual role, behaving in both the expected manner, and in

[24] C. Harries, and W. Nagel, Kolloid Zeits. **33**, 248 (1923).
[25] T. H. Barry, *Natural Varnish Resins* (Ernest Benn Ltd., London, 1932), p. 265.
[26] W. P. Henckel, Polytechnic Inst. of Brooklyn, Bachelor of Chemical Engineering Thesis, 1934.
[27] S. Ranganathan, and R. W. Aldis, Indian Lac Research Institute, Research Note No. 13 (1933).
[28] R. W. Aldis, Indian Lac Research Institute, Research Note No. 18 (1934).
[29] W. H. Gardner, British Plastics **6**, 524 (1935).

functioning as a catalyst to promote ester or anhydride formation between the lac molecules themselves. This and the large accelerated effects with mineral acids are easy to explain, since it is well known that acids of low *pH* are catalysts for ester- and anhydride-forming reactions.

In contrast, it is not so simple to account for the reactions between shellac and ketones or aldehydes. Our lack of knowledge concerning the constitution of the possible benzoid type of hydroxy-acids found in these resins permits only surmise. There has been no evidence as yet obtained to show that lac contains phenolic, ketone or unsaturated groups capable of causing condensation with these reagents. One would be misled by attempting to draw conclusions from analogy to other natural resins.

Effect of Water

It is well known that moisture and pressure have a very marked effect upon the plastic properties of lac. The presence of free water in lac definitely lowers the softening point and increases the fluidity. This has been one of the reasons why different chemists have been unable to obtain concordant results for the fluidity of the same samples of lac, and different methods carried out under similar conditions give varying values.

The elimination of occluded moisture from a test sample, whether it be ground lac or a film, is not as simple as it would appear. Drying at any temperature from the point of view of a high degree of precision does not preclude the possibility of the splitting off of water as a result of chemical condensation. A fairly representative value of occluded moisture however can be obtained by drying overnight at 41–43°C.[30]

Although water is not a solvent for shellac, this resin is so highly hydroxylated that it is readily wet by this and similar liquids. Rules generally enumerated for solubility can in this case be applied to adsorption. The strong tendency to retain both water and other liquids suggests that they may be incorporated in the intermicellular spaces in much the same manner as that of a solid solution. The removal of the ether soluble resins tends to reduce susceptibility to water.[31]

The reason that pressure retards polymerization[19] is probably due to the fact that water as one of the reaction products is unable to escape, since it has been demonstrated that heat polymerization involves more or less reversible reactions. In the depolymerization the only marked difference, which has been noted, between the original lac and lac regenerated from heat hardened material is a greater tendency on the part of the latter to repolymerize or age.[32] This tendency to polymerize may be due more to a greater tendency to form aggregates than to condense.

Small amounts of alcohol vapors are also adsorbed by shellac in much the same manner as those of water, and have a similar effect in increasing the fluidity. There have been instances where machine-made lacs manufactured by a solvent process have been rejected because of a high fluidity, which, after storage and loss of residual solvent, could then meet the individual consumer's specifications.

Methods Used for Determining Fluidity

Industry has employed two distinct types of apparatus based upon the same general principle of natural flow at a given temperature for measuring relative fluidities of different lacs. One consists of an electrically heated block containing a well for melting the resin, and a grooved channel calibrated in inches for measuring the flow when the block is tilted at a 15° angle from the horizontal. The results are expressed as the length of time required for the lac to reach each mark at any selected temperature. The other consists of a glass tube so mounted in a water, steam or oil bath that the shellac can be melted with the tube in a vertical position, and the degree of flow determined at a position 15° from the horizontal. In this case, the length of flow for an agreed length of time at a selected temperature is recorded. Unfortunately none of these methods give results which can be expressed in rheological units, but they do give a simple, rapid qualitative method for measuring the relative effects of different factors upon fluidity.

[30] J. W. Paisley, F. S. White, B. H. Knight, G. E. Ashby, R. V. Townend, H. E. Wolff and W. H. Gardner, *Official Methods of Analysis, Specifications and General Information on Shellac and Bleached Shellac* (U. S. Shellac Importers Assoc., New York, 1934), p. 15.

[31] L. C. Verman, J. Soc. Chem. Ind. **54**, 85 (1935).

[32] M. Rangaswami, and R. W. Aldis, Indian Lac Research Institute, Bull. No. 19 (1934).

The McLain metal block has been patented by the Westinghouse Electric and Manufacturing Company in the United States[33] and by the Metropolitan Vickers in England[34] and has been recommended by the British National Committee of the International Electrotechnical Commission.[35]

A committee of the American Society of Testing Materials has also considered this method but has been unable to obtain consistent results between different laboratories, even where the individual laboratories have used it for some years as a control. The difficulty may in part be due to differences in the conditioning of the samples, but more likely to heat variations in the channel in the block during the period of flow and to uncontrolled atmospheric conditions surrounding the apparatus during test.[36] Each laboratory in this country has constructed its own apparatus which may also account for some of the variations. However, at best, this instrument is costly, as compared with the glass tube in the other method.

The alternate method, sometimes called the V-tube method,[25] was originally developed by J. W. Paisley of the Rogers-Pyatt Shellac Company. The General Electric Company, the Victor Talking Machine Company, Angelo Brothers Ltd., and the Bell Telephone Laboratories have all developed modifications from the original apparatus. This method is now being studied by the American Society of Testing Materials with a view to its adoption as an American standard.[37] With the Bell Telephone Laboratories modification it should be possible by changes in technique to obtain data comparable with that obtained by the McLain block test. From a commercial standpoint, this is desirable since relative fluidity could then be readily correlated with length of "life" or polymerization under heat for general industrial purposes.

[33] J. R. McLain, U. S. Patent 163785, Aug. (1927).
[34] British Patent 254239, July (1928).
[35] E. J. Parry, *Shellac, Its Production, Manufacture, Chemistry, Analysis, Commerce and Uses* (Sir Isaac Pitman & Sons, Ltd., London, 1935), p. 114.
[36] B. Gross and W. H. Gardner, unpublished investigations.
[37] C. F. Hanson, Proc. A. S. T. M. **34**, 502 (1934).

A Method of Analyzing Experimental Results Obtained from Elasto-Viscous Bodies

ANDREW GEMANT, *Engineering Laboratory, Oxford University, Oxford, England*

(Received February 27, 1936)

An attempt is made to develop a general method of analyzing experimental results concerned with the behavior of elasto-viscous bodies. It is shown that the transient equation for relaxation processes can be deduced from the experimental impedance formula, and that relaxation experiments are in agreement with the deduced transient equation. A differential equation relating stress, strain and time is derived from the experimental impedance formula. The relation of this equation toithe well-known Maxwell equation is discussed.

INTRODUCTION

THIS paper is concerned with a method of analyzing quantitative experimental results obtained from liquids showing structure-viscosity, or, in more general terms, from elasto-viscous bodies.

Maxwell established a differential equation relating stress, strain and time for a volume element of an elasto-viscous body subjected to external forces. Attempts have been made to apply this equation to quantitative experiments involving either relaxing or steady vibrational disturbances, and to deduce the shape of a transient for the former and the expression for the impedance as a function of frequency for the latter. Both kinds of processes have recently been investigated experimentally, and a definite deviation has been found between the quantitative results and those predicted by means of the Maxwell equation.

In the present paper an attempt is made to clear up this difficulty. In the first part of the paper a method (based on the Heaviside calculus) is developed, by means of which relaxation and oscillation experiments can be related. We commence with a formula for the impedance which can be derived experimentally; and from it deduce directly the transient equation for a sudden disturbance. Later it will be shown that some experimental relaxation data supports these deductions in a semiquantitative manner. In the second part we shall proceed to derive the differential equation between stress, strain and time from the experimental formulae for the impedance; and, indeed, this differential equation will prove different from the Maxwell relation. This derived differential equation is not, however, of a fundamental type valid for any volume element, as is the Maxwell relation, but it represents a kind of "equivalent equation" which is valid only for the body as a whole. This equation, which is not necessarily a contradiction of the Maxwell relation, is useful in that the experiments, whether of the relaxation or oscillation type, yield directly only the "equivalent" form shown in this paper.

I. The transient phenomenon

Let us first summarize the results of Maxwell's theory. Having an elasto-viscous body, the following differential equation is assumed to hold for each volume element:

$$dS/dt = F/\eta + (1/E)(dF/dt), \tag{1}$$

where S = strain, F = stress, η = viscosity, E = shear modulus of elasticity and t = time. Applying it to a transient process, the result will be

$$F = \mathbf{S}_0\eta(1-e^{-t/\tau}). \tag{2}$$

Here $\mathbf{S}$ stands for dS/dt and $\mathbf{S}_0$ denotes a *sudden* deformation velocity of constant value, applied from $t=0$ onwards. For $t=0$ the stress must be evidently zero, whereas for $t=\infty$ it is simply given by the viscosity. The transition occurs exponentially with the time constant τ, the latter being

$$\tau = \eta/E. \tag{3}$$

If, on the other hand, a sinusoidal oscillation with frequency ν is applied to the system, a certain elasto-viscous resistance will be exerted by the system, which can be written as a complex impedance. Using the latter in its most general meaning, suggested by the author (for molecular processes also), we denote it by V, as complex

viscosity. Assuming Eq. (1) to be correct, we obtain for V

$$V=\eta(1-i\omega\tau)/(1+(\omega\tau)^2), \tag{4}$$

where $\omega=2\pi\nu$.

Two ways are possible for checking this formula for the impedance. One is to investigate dynamic phenomena connected with the macroscopic flow in elasto-viscous media. This method has been followed up recently by W. Philippoff.[1] The other method is, to extend the conception of an elasto-viscous resistance to molecular processes, as has been done by the author[2, 3] in introducing the conception of a complex viscosity. As a result of this extension, dielectric losses on the one hand, damping of acoustical waves on the other, both in amorphous media, can be utilized from this standpoint, in that they can be related to the above-mentioned impedance, thus allowing an experimental check of the theoretical formula.

Philippoff investigated the real part of (4), and found that the exponent of the argument $(\omega\tau)$ is not 2, but much nearer to unity (approximately $\frac{3}{4}$). The author's discussion of dielectric losses in glassy materials indicates too that an exponent of 1 instead of 2 is more in keeping with observations. So far, nothing quantitative is known about the exponent of $(\omega\tau)$ in the numerator. All in all, the general form of (4) seems to be right, except that the two exponents have to assume other values.

The present question is to determine the shape of the transient when the expression of the impedance is known. Such problems have been treated in a more or less general way by several authors, for instance R. M. Foster[4] and W. Cauer.[5]

We shall try to solve our special problem by means of the Heaviside calculus.[6] One has to be careful with the latter not to apply it to cases in which it may lead to wrong results. The method we adopted here, is the following. First we shall show how the Heaviside calculus works for Eq. (4), leading directly to Eq. (2). We then shall proceed with a possibly small deviation from Eq. (4) and find the corresponding equation for Eq. (2). If the result is right, it must differ from Eq. (2) only to a little extent and in a physically reasonable manner. We shall then try to generalize the rule found in this way and calculate the case which is suggested by experiment.

In applying the method of Heaviside to Eq. (4), one has to consider that according to the definition of V, the stress F is given for a sinusoidal $\mathbf{S}$ by

$$F=V\mathbf{S}. \tag{5}$$

Heaviside applies the same formula for the transient, in substituting the operator p instead of $i\omega$. Thus

$$F=\mathbf{S}_0\eta(1-p\tau)/(1-(p\tau)^2). \tag{6}$$

This is really the symbolic solution for the transient Eq. (2); it has only to be transformed according to Heaviside's rules. We have to expand the operator-equation into a power series with descending powers of p:

$$F=\mathbf{S}_0\eta(1/p\tau-1/(p\tau)^2+1/(p\tau)^3-\cdots) \tag{7}$$

and replace $1/p^n$ by $t^n/n!$:

$$F=\mathbf{S}_0\eta\left(\frac{t}{\tau}-\frac{1}{2!}\left(\frac{t}{\tau}\right)^2+\frac{1}{3!}\left(\frac{t}{\tau}\right)^3-\cdots\right), \tag{8}$$

which is exactly Eq. (2).

As already mentioned, the general form of Eq. (4) has to be preserved only changing the exponents:

$$V=\eta(1-i(\omega\tau)^g)/(1+(\omega\tau)^h), \tag{9}$$

where g and h are any positive numbers.

In order to keep as close as possible to the original calculation, we must be able to transform Eq. (9) into

$$V=\eta(1-(p\tau)^g)/(1-(p\tau)^h), \tag{10}$$

which is analogous to Eq. (6). Since $p=i\omega$, Eq. (10) is equivalent to

$$V=\eta(1-(i\omega\tau)^g)/(1-(i\omega\tau)^h). \tag{11}$$

Comparing Eq. (9) with Eq. (11), we have the conditional equations $i^g=i$ and $i^h=-1$. This again means that

$$g=(4j\pm1)/(4k\pm1) \tag{12}$$

[1] W. Philippoff, Physik. Zeits. **35**, 884 (1934).
[2] A. Gemant, Trans. Faraday Soc. **31**, 1582 (1935).
[3] A. Gemant, Physics **6**, 363 (1935).
[4] R. M. Foster, Bell Sys. Tech. J. **3**, 259 (1924).
[5] W. Cauer, Archiv f. Elektrotechnik **17**, 355 (1926).
[6] J. R. Carson, *Electric Circuit Theory and Operational Calculus* (McGraw-Hill, New York, 1926).

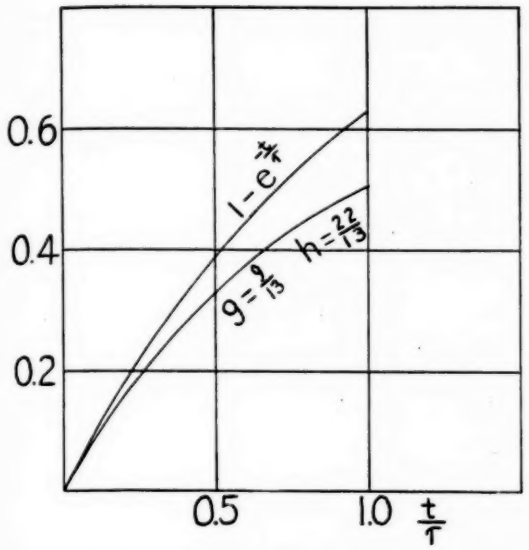

FIG. 1. Transient corresponding to impedance: $V=\eta\frac{1-i(\omega\tau)^g}{1+(\omega\tau)^h}$. (Exponential for comparison.)

and $$h=(4l+2)/(2m+1), \qquad (12a)$$

where j, k, l, m = any positive integers.

Let us choose two numbers for g and h, fulfilling Eqs. (12) and (12a) and at the same time not being too different from 1 and 2 (their values in Eq. (4)). $g=9/13$ and $h=22/13$ will fulfill this condition. Putting them into Eq. (10) and expanding in a series corresponding to Eq. (7), we obtain

$$F=\mathbf{S}_0\eta\left(\frac{1}{p\tau}-\frac{1}{(p\tau)^{22/13}}+\frac{1}{(p\tau)^{35/13}}-\frac{1}{(p\tau)^{44/13}}+\cdots\right). \qquad (13)$$

The question now arises whether Heaviside's rule $1/p^n\rightarrow t^n/n!$ is valid for n being not an integer but a positive fraction. This can be answered in the affirmative, as can be seen for instance by the integral equation Carson makes use of to justify Heaviside's power series method. This equation is

$$\int_0^\infty e^{-pt}t^n dt=n!/p^{n+1}.$$

Now this equation holds for any positive value of n, not only for integers.[7] Thus we write for Eq. (13)

$$F=\mathbf{S}_0\eta\left(\frac{t}{\tau}-\frac{1}{(22/13)!}\left(\frac{t}{\tau}\right)^{22/13}+\frac{1}{(35/13)!}\left(\frac{t}{\tau}\right)^{35/13}-\cdots\right). \qquad (14)$$

Here the factorials as applied to fractions are those known as Legendre or Gauss functions. If the result Eq. (14) is likely to be correct, it must be nearly an exponential. Since the powers g and h, indicating the variation with frequency of the impedance, are less than 1 and 2, the transient may be expected to run more flat than the exponential. Fig. 1 shows the values in the brackets in both Eq. (2) and Eq. (14) for the range $t/\tau=0$ to 1, in agreement with our predictions.

Considering the transient, which belongs to the impedance (9) with the exponents 9/13 and 22/13 to be given by (14), the rule expressed by Eqs. (10), (13) and (14) will now be generalized independently, whether Eq. (12) is fulfilled or not. To justify this generalization, two points should be mentioned.

(a) Even if a fraction (say 2/3) does not fulfill Eq. (12), we may *approach* it with any desired degree of accuracy, by multiplying the nominator and denominator by n, where n is a large integer

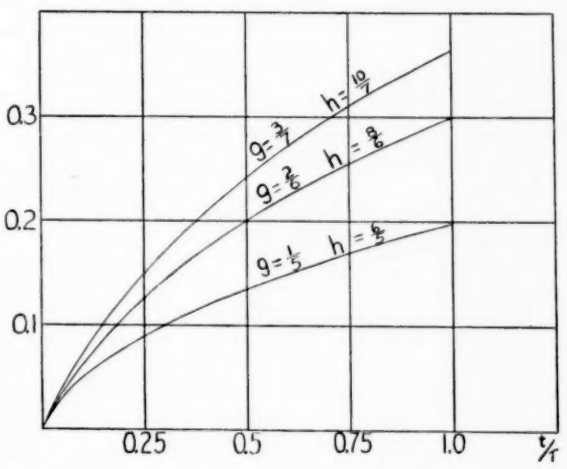

FIG. 2. Transients corresponding to impedance: $V=\eta\frac{1-i(\omega\tau)^g}{1+(\omega\tau)^h}$.

[7] E. Jahnke and F. Emde, *Tables of Functions* (Leipzig, 1933), p. 94.

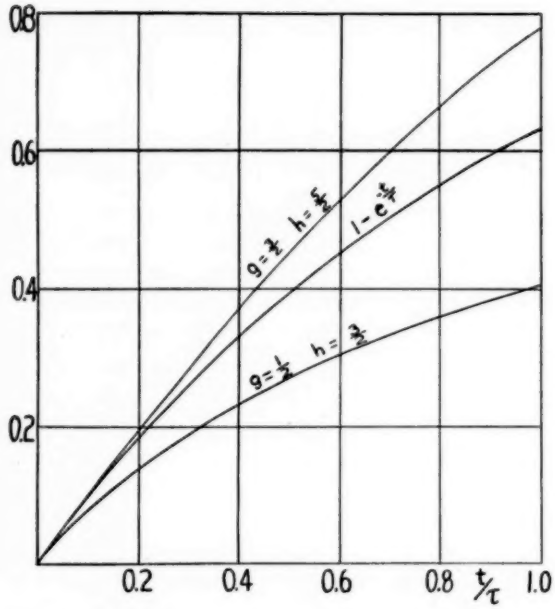

FIG. 3. Transients corresponding to impedance:

$$V=\eta\frac{1-i(\omega\tau)^{g}}{1+(\omega\tau)^{h}}.$$ (Exponential for comparison.)

and adding 1 to each (with $n=1000$, for instance 2001/3001). The new fraction fulfills Eq. (12) exactly and by increasing n beyond any limit can be made to coincide with the original fraction. In other words, the rule leading to the above solution is really independent of Eq. (12).

(b) We calculate the transient belonging to the impedance with $g=3/7$ and $h=10/7$, fulfilling Eq. (12). The curve is shown in Fig. 2. Further, we calculate the case $g=1/5$, $h=6/5$, again fulfilling (12); the result is shown in Fig. 2. Now we calculate the intermediate case $g=2/6$, $h=8/6$, *not* obeying Eq. (12). However, the curve as shown in Fig. 2 lies exactly between the other two, showing the correctness of the method.

In Fig. 3 the curve in the middle is the exponential. The lower curve corresponds to the values $g=1/2$, $h=3/2$, the upper curve to the values $g=3/2$, $h=5/2$. The latter ones are bigger than those belonging to Eq. (4), namely, 1 and 2, indicating a stronger variation with frequency. Correspondingly, the transient must be *steeper* than the exponential, in agreement with the computation.

The above series do not allow us to foretell how all these curves behave for higher values of t/τ; this would require an asymptotic expansion.

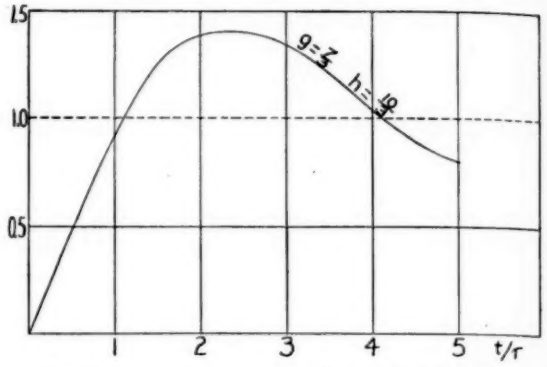

FIG. 4. Transient corresponding to impedance:

$$V=\eta\frac{1-i(\omega\tau)^{g}}{1+(\omega\tau)^{h}}.$$

Purely physical reasons suggest that they all ought to reach unity for $t/\tau=\infty$. If, however, the curve is steeper than the exponential, the sum of the series (which is always convergent) reaches values higher than 1. If the result is still correct, a kind of oscillatory function is to be expected. To indicate the correctness of this assumption, the case $g=7/3$, $h=10/3$ has been calculated for arguments as high as 5. The result is shown in Fig. 4, fully justifying the predictions. This curve corresponds to the series:

$$F=\mathbf{S}_0\eta\left(\frac{t}{\tau}-\frac{1}{(10/3)!}\left(\frac{t}{\tau}\right)^{10/3}+\frac{1}{(13/3)!}\left(\frac{t}{\tau}\right)^{13/3}-\frac{1}{(20/3)!}\left(\frac{t}{\tau}\right)^{20/3}+\cdots\right). \quad (15)$$

In general terms, starting with Eq. (9), the result will be given by

$$F=\mathbf{S}_0\eta\left(\frac{1}{(h-g)!}\left(\frac{t}{\tau}\right)^{h-g}-\frac{1}{h!}\left(\frac{t}{\tau}\right)^{h}+\frac{1}{(2h-g)!}\left(\frac{t}{\tau}\right)^{2h-g}-\frac{1}{(2h)!}\left(\frac{t}{\tau}\right)^{2h}+\cdots\right). \quad (16)$$

Eqs. (9) and (16) are the general results of this investigation. As to the values of g and h, the only restriction is that they must be positive and that

$$g<h. \quad (17)$$

This follows from the physical requirement that the resistance must ultimately fall with increasing frequency.

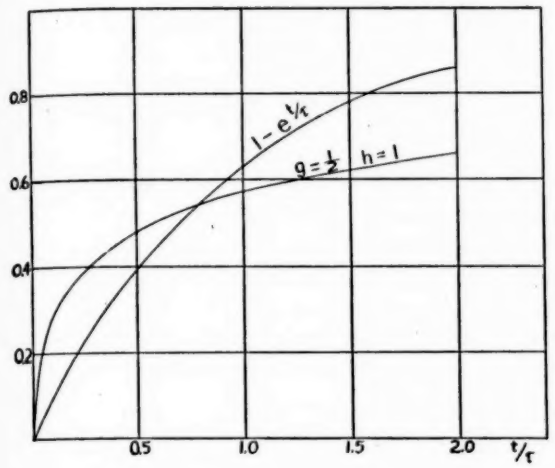

FIG. 5. Transient corresponding to experimental impedance-formula: $V=\eta\frac{1-i(\omega\tau)^{0.5}}{1+(\omega\tau)}$

Having given already several numerical examples of Eq. (16), one more should be added which seems to be the nearest to *experimental* results concerning impedance. As already stated, h is definitely nearer to unity than to 2. Let us assume, therefore, that $h=1$. No indication exists as to the real value of g, but since in the Maxwell assumption $h=2g$, let us assume that $g=\frac{1}{2}$. We start with the complex viscosity

$$V=(1-i(\omega\tau)^{0.5})/(1+\omega\tau) \qquad (18)$$

as the best in agreement with experiments on steady vibrations, and have therefore

$$F=S_0\eta\left(\frac{1}{0.5!}\left(\frac{t}{\tau}\right)^{0.5}-\frac{t}{\tau}+\frac{1}{1.5!}\left(\frac{t}{\tau}\right)^{1.5}-\frac{1}{2}\left(\frac{t}{\tau}\right)^{2}+\cdots\right) \qquad (19)$$

as the transient equation deduced from Eq. (18). The curve for Eq. (19) is indicated in Fig. 5 up to $t/\tau=2$. The exponential is shown for comparison. The difference is striking: the graph for Eq. (19) is steeper at the beginning and later more flat than the exponential; they cross at the ordinate 0.54.

A rigorous experimental verification of our theory would only be possible if both oscillation and relaxation experiments were carried out on the same process.[8] The author is not aware of any such experiments. There is, however, a very careful investigation by Schofield and Scott Blair[9] on the relaxation of flour dough, which is a typical elasto-viscous medium. They too noticed that the relaxation curve is not an exponential. In Fig. 6 we reproduce a typical curve shown on p. 714 of the paper cited above, together with an exponential, the latter chosen in a way as to cross the first at the ordinate 0.54. There is a striking similarity between Figs. 5 and 6 which of course cannot be more than a semi-quantitative one, since the impedance for flour dough will not be identical with Eq. (18). However, there cannot be expected more, the quantitative data available being so scanty.

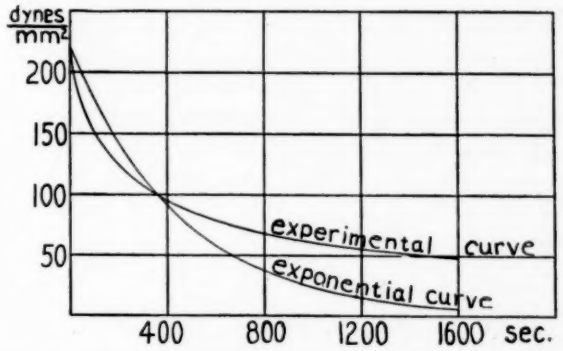

FIG. 6. Relaxation curve obtained with flour dough by Schofield and Scott Blair. (Exponential for comparison.)

II. The differential equation

The Heaviside calculus not only establishes a relation between the impedance and the transient, but also one between the impedance and the underlying differential equation. The first of these is governed by the substitution $p=i\omega$, the second, with which we are now concerned, by $p=d/dt$.

Following up the method used in the previous section, it should be indicated first, how the method works in the case of the Maxwell relation. The impedance Eq. (4) leads us directly to the operator Eq. (6). Now Eq. (6) may be written in the form

$$F=S_0\eta/(1+p\tau), \qquad (20)$$

[8] It appears that g and h are probably independent of the material, but depend on the process under investigation.

[9] R. K. Schofield and G. W. Scott Blair, Proc. Roy. Soc. **A138**, 707 (1932).

or slightly rearranged and in considering Eq. (3)

$$\mathbf{S}_0 = F/\eta + pF/E.$$

This equation leads to the differential form by replacing the constant $\mathbf{S}_0$ by the more general $\mathbf{S}$ and p by d/dt:

$$\mathbf{S} = F/\eta + (1/E)dF/dt.$$

This is exactly the Maxwell relation.

In following this way of substitution, we shall find the differential equation belonging to the general impedance (9). It will be simpler, however, to consider more in detail Eq. (18) instead of Eq. (9), since its particular form is in keeping with experimental observations. The method, according to which the more general case (9) has to be treated, will be indicated briefly.

Eq. (18) leads to the operational equation, analogous to Eq. (20),

$$F = \mathbf{S}_0\eta/(1+(p\tau)^{0.5}), \tag{21}$$

or rearranged

$$\mathbf{S}_0 = F/\eta + (p/\eta E)^{\frac{1}{2}}F. \tag{22}$$

In going over to the differential form, the question arises as to the meaning of $p^{\frac{1}{2}}$. In accordance with Heaviside,[10] we have to identify $p^{\frac{1}{2}}$ with $d^{\frac{1}{2}}/dt^{\frac{1}{2}}$. In other words, so-called fractional differentials have to be introduced. Our differential equation will read:

$$\mathbf{S} = F/\eta + 1/(\eta E)^{\frac{1}{2}}(d^{\frac{1}{2}}F/dt^{\frac{1}{2}}). \tag{23}$$

The next step is to prove that Eq. (19), as a transient, fulfills the differential Eq. (23). But in order to show this, we have to know, how to calculate a half-differential. Quite generally, having a function $y=hx^n$ the qth differential quotient, where q is any number, integer or fraction, is given by:

$$y^{(q)} = h(n!/(n-q)!)x^{n-q} \tag{24}$$

and correspondingly the half-differential by:

$$d^{\frac{1}{2}}y/dx^{\frac{1}{2}} = h(n!/(n-0.5)!)x^{n-0.5}. \tag{24a}$$

Performing this operation with each term of (19), we get

$$\frac{1}{(\eta E)^{\frac{1}{2}}}\frac{d^{\frac{1}{2}}F}{dt^{\frac{1}{2}}} = \mathbf{S}_0\left(1 - \frac{1}{0.5!}\left(\frac{t}{\tau}\right)^{0.5} + \frac{t}{\tau} - \frac{1}{1.5!}\left(\frac{t}{\tau}\right)^{1.5} + \cdots\right).$$

Adding this to F/η, the result is $\mathbf{S}_0$, hence fulfilling Eq. (23).

It is therefore justifiable to consider Eq. (23) as one of the most probable differential relations which this analysis leads to, instead of the Maxwell equation.

One could object that $d^{\frac{1}{2}}y/dx^{\frac{1}{2}}$ has a mathematical meaning at the best, but no physical meaning whatever; so it seems to be unlikely to have a fractional differential in a basic differential equation, controlling the behavior of any real medium. In order to eliminate this difficulty, I should like to point out briefly a certain analogy between our problem and that of the non-inductive cable which can be treated by the Heaviside calculus. The differential equations at any point of the cable (resistance R and capacitance C per unit length) contain the first differential quotients of voltage U and current I with respect to cable axis and time. But taking the *origin* of an infinite cable, and establishing the relation between U and I at the origin, Heaviside obtains the operational equation

$$I = (Cp/R)^{\frac{1}{2}}U.$$

But this is nothing else than a differential equation with half-differential

$$I = (C/R)^{\frac{1}{2}}d^{\frac{1}{2}}U/dt^{\frac{1}{2}}, \tag{25}$$

and we see here that such equations really may occur in a physically realizable case (a finite cable shunted at the end by its characteristic resistance is equivalent to an infinite cable).

That Eq. (25) is correct, can be seen by considering a definite case. Let us have $U=U_0$, a constant. It is known that in that case

$$I = (C/R\pi t)^{\frac{1}{2}}U_0. \tag{26}$$

But applying Eq. (24) to a constant U_0, the result will be exactly Eq. (26), considering that $(-0.5)! = \pi^{\frac{1}{2}}$.

The conclusion from this analogy is the following: although the *primary* differential equations do not contain fractional differentials, it is still

[10] O. Heaviside, *Electromagnetic Theory*, Volume II (London, 1922).

possible to derive certain secondary differential equations containing fractional differentials. It occurs in the case of the cable, if we take the values only at the origin; considering, so to speak, the equivalent electric circuit to the infinite cable. The same might hold for our problem. The measurements never refer to one volume element but always to a more or less large system, revealing the value of the impedance occurring at certain parts of the system as a resultant of the whole. We cannot expect that the underlying differential equation will be that of the volume element itself; what we obtain is a kind of "equivalent differential equation" belonging to an equivalent unit, which replaces the more or less extended real system. That is the reason why all the experiments indicate a basic equation of the type of Eq. (23) instead of that belonging to the volume element suggested by Maxwell. This latter may be deduced only by means of a more detailed analysis from the former; it is very likely of the Maxwellian type, if not identical with it.[11]

We close this part by showing how the differential equation may be deduced from the general form (9) of the impedance. Eq. (10) can be brought into the form

$$\mathbf{S}_0 = F(1-(p\tau)^h)/\eta(1-(p\tau)^g),$$

and the fraction has to be expanded into a series. If h/g is an integer, the series is finite and given by

$$1+(p\tau)^g+(p\tau)^{2g}+\cdots(p\tau)^{h-g},$$

and the differential equation contains several terms, namely

$$\mathbf{S}=\frac{F}{\eta}+\frac{\tau^g}{\eta}\frac{d^gF}{dt^g}+\frac{\tau^{2g}}{\eta}\frac{d^{2g}F}{dt^{2g}}+\cdots\frac{\dot{\tau}^{(h-g)}}{\eta}\frac{d^{(h-g)}F}{dt^{(h-g)}}. \quad (27)$$

Fortunately Eq. (18), which seems to be the best approximation to experiments, leads to an equation with only one differential term.

---

[11] This is not the only possible physical interpretation. It might well be that a consideration of several coexisting time-constants, due to inhomogeneity of the material, will lead to an equivalent differential equation, similar to (23).